



HAL
open science

Contributions aux réseaux intelligents - The Next Generation Networking and Internet

Sami Souihi

► **To cite this version:**

Sami Souihi. Contributions aux réseaux intelligents - The Next Generation Networking and Internet. Informatique [cs]. Université Paris-Est, 2023. tel-04344283

HAL Id: tel-04344283

<https://hal.u-pec.fr/tel-04344283v1>

Submitted on 14 Dec 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

École doctorale : Mathématiques et STIC

L'Université Paris Est Créteil

Manuscrit présenté par

Sami SOUIHI

le 17 Février 2023

en vue de l'obtention d'une :

Habilitation à Diriger des Recherches

Contributions aux réseaux intelligents
The Next Generation Networking and Internet

Jury

Isabelle GUERIN LASSOUS,	Professeur	Rapporteur
Thierry DIVOUX,	Professeur	Rapporteur
Stéphane MAAG,	Professeur	Rapporteur
Christophe CHASSOT,	Professeur	Examineur
Olivier FESTOR,	Professeur	Examineur
Gerardo RUBINO,	Professeur	Examineur
Abdelhamid MELLOUK,	Professeur	Examineur

Résumé

La croissance continue du trafic de données, l'émergence de la virtualisation des réseaux ainsi que l'utilisation sans cesse croissante d'équipements mobiles dans les réseaux modernes ont mis en lumière les nombreux problèmes inhérents à l'architecture conventionnelle de l'Internet. Ainsi, la tâche de gestion et de contrôle des informations devient de plus en plus complexe et spécialisée. Dans ce contexte, il est devenu urgent de repenser la philosophie et l'architecture des mécanismes de contrôle des réseaux en rendant ces derniers plus autonomes avec une adaptabilité accrue à leurs contextes dynamiques. Pour concevoir ces nouveaux réseaux, plusieurs stratégies de recherche ont été proposées afin de relever les défis de l'évolutivité, de la fiabilité et de la disponibilité du trafic en temps réel, garantissant ainsi la qualité de l'expérience de l'utilisateur. Cette thèse d'Habilitation à Diriger des Recherches résume mes travaux de recherche effectués dans ces thématiques au cours de ces dernières années (2013-2022). Elle décrit mes principales contributions dans ce qui est appelé aujourd'hui « les réseaux intelligents ou encore les réseaux de nouvelle génération ou pour certains les réseaux adaptatifs ». Elle met en avant également ma méthode et ma philosophie de travail, fruit de mon expérience de ces dernières années ainsi que certaines perspectives pour des travaux futurs.

Abstract

The continuous growth of data traffic, the emergence of network virtualization and the ever-increasing use of mobile devices in modern networks have highlighted the various problems associated with the conventional Internet architecture. Therefore, monitoring and controlling tasks are becoming increasingly complex and specialized. In this context, it has become urgent to rethink the philosophy and architecture of network control mechanisms by making them more autonomous and adaptive to the dynamic contexts. To design these new networks, several research strategies have been proposed to address the scalability, reliability and availability challenges of real-time traffic, thus ensuring the quality of the user experience. This Habilitation thesis summarizes my research work in these areas over the last few years (2013-2022). It describes my main contributions in what is now called "new generation networks or adaptive networks". It also presents my work methods and philosophy, based on these last years' experience, as well as some perspectives for future works.

Table des matières

Table des matières	vii
Liste des figures	ix
Liste des tableaux	xi
Préambule	1
1 Introduction Générale	3
1.1 Brève histoire des réseaux intelligents	4
1.2 Sélection de mes activités de recherches	4
1.3 Organisation du manuscrit	6
2 Curriculum Vitae	9
2.1 Présentation synthétique des thématiques de recherche	10
2.2 Encadrement doctoral et scientifique	10
2.3 Diffusion, rayonnement et vulgarisation	12
2.4 Responsabilités scientifiques	18
2.5 Responsabilités au niveau local	20
2.6 Participation aux conseils des composantes	20
3 Le paradigme SDN, le chaînant manquant	23
3.1 Introduction et contexte	24
3.2 Motivation et état de l'art	24
3.3 contribution 1 : placement des contrôleurs	27
3.4 contribution 2 : modèle de cohérence continue	35
3.5 Validation expérimentale	37
3.6 contribution 3 : modèle de cohérence adaptatif	39
3.7 Conclusion et travail réalisé	45
3.8 Synthèses	47
4 Supervision du réseau : de la collecte à l'élaboration des connaissances	51
4.1 Introduction et contexte	52
4.2 Problématique : Collecte des données et élaboration des connaissances	52
4.3 Contribution 1 : Analyse et classification du trafic	52
4.4 Contribution 2 : Estimation de la QoE	61
4.5 Conclusion et travail réalisé	68
4.6 Synthèses	69

5	Contrôle adaptatif du réseau : Le Réseau de demain, un réseau autonome	75
5.1	Introduction et contexte	76
5.2	Problématique : qu'est ce qu'un réseau autonome et comment le mettre en oeuvre	76
5.3	Contribution 1 : Routage adaptatif ou comment auto-corriger les défaillances	77
5.4	Mécanisme adaptatif de routage de segments	79
5.5	Résultats expérimentaux	81
5.6	Contribution 2 : La diffusion vidéo à débit variable (ABR) sensible à la QoE : Un exemple auto-optimisation	83
5.7	Évaluation	88
5.8	Synthèses	90
6	Conclusion Générale	97
6.1	Contributions et retours d'expérience	98
6.2	Autres travaux	98
6.3	Méthode de travail : organisation agile pour la recherche	100
6.4	Défis scientifiques et perspectives	102
7	Articles co-rédigés	105

Liste des figures

1.1	Nuage de mots créé à partir du contenu de toutes mes publications post-thèse	6
3.1	Réseaux conventionnels vs SDN	24
3.2	Architecture des réseaux SDN	25
3.3	Évolution de l'état de l'art sur les SDNs (en nombre de publications)	25
3.4	Stratégie 1, 2 et 3 : Mesures de performance basées sur la latence	32
3.5	Strategie 4 : Mesures de fiabilité : (latences maximales dans les scénarios sans défaillance)	33
3.6	Strategie 4 : Mesures de la performance	34
3.7	Scénario 1 : Trafic inter-contrôleur capturé (en paquets par seconde) pendant la période du scénario de test (à l'aide de Wireshark)	38
3.8	Gain en surcharge anti-entropie de l'ONOS-WAC par rapport à l'ONOS par rapport à ONOS en fonction du nombre de contrôleurs dans le cluster	39
3.9	Algorithme Q-Learning.	40
3.10	Transition d'état des commutateurs et des liens	42
3.11	ReadDelay et WriteDelay avec les valeurs fixes de $N_W = 18$ et $N_R = 1$.	43
3.12	ReadDelay et WriteDelay avec différentes valeurs de ThR	44
3.13	Version staleness et time staleness avec différentes valeurs de ThR	44
3.14	Survol avec différentes valeurs de ThR	45
4.1	Structure générale	58
4.2	Multi-Classifier	59
4.3	Précision macro-moyenne, rappel macro-moyenne et f1-score macro-moyenne dans différents ensembles de données	59
4.4	Macro-moyenne de la précision, macro-moyenne du rappel et macro-moyenne du f1-score pour différentes fonctions de perte	60
4.5	Précision, Recall et F1-score de notre méthode proposée sur cinq types de services basés sur QUIC	60
4.6	Corrélation de la QoE avec les paramètres vidéo	63
4.7	Distribution de la meilleure méthode de prédiction pour chaque échantillon	64
4.8	Présentation générale du modèle de prédiction Belief Functions Combination (BFC)	64
4.9	Description du modèle de croyances transférable (TBM)	65
4.10	Processus d'estimation de la qualité d'expérience	67
4.11	Performance de prédiction en termes de taux de MAE lorsque la taille des données d'apprentissage varie.	68
4.12	Performance de prédiction en termes de taux de MAE	68
5.1	Le cadre du mécanisme SR adaptatif basé sur le SDN.	79
5.2	Le mécanisme SR basé sur RL	80

5.3	Score MOS moyen et écart type des trois algorithmes de sélection.	82
5.4	Le score MOS de trois mécanismes SR.	83
5.5	Architecture de la méthode de streaming à débit binaire adaptatif (ABR) proposée	84
5.6	Description de l'agentABR	85
5.7	Description du système RL	86
5.8	Exécution de l'approche Q2ABR	87
5.9	Processus de détection du point de rupture	87
5.10	Résultats de la comparaison des approches ABR en termes de nombre de rebufferades	88
5.11	Résultats de la comparaison des méthodes ABR en termes de nombre de changements de qualité n_{sw}	89
5.12	Résultats de la comparaison des méthodes de RAP en termes de données totales téléchargées (Kbits)	89
5.13	Résultats de la comparaison des méthodes RAB en termes de moyenne et de variation des scores MOS prédits	90
6.1	Implémentation de SINA	99

Liste des tableaux

3.1	Principales caractéristiques des contrôleurs SDN existants	26
4.1	Précision, rappel et score F1 pour la première étape de la classification à l'aide de l'algorithme de forêt aléatoire	57
4.2	Tableau de comparaison des méthodes testées	63

Préambule

« *Coincidence is God's way of remaining Unknown Author.* »

Albert Einstein

Jeune, j'avais peu de certitudes sur mon devenir professionnel. Je savais que je voulais devenir ingénieur en informatique, sans doute grâce à mon père mais je n'ai jamais eu la volonté et l'envie de devenir enseignant-chercheur.

...Et pourtant ...

Les circonstances, mes rencontres, mes lectures et notamment Celle de David D. Clark¹, m'ont fait réfléchir à la possibilité d'horizons nouveaux. Je me suis inscrit en Master 2 et je me suis mis à la recherche d'un cadre où je pouvais travailler sur une thématique qui me passionnait. Cette thématique faisait rêver mon "moi enfant" et satisfaisait mon "moi ingénieur". Dans cette démarche, j'ai eu la chance de trouver en la personne du Professeur Abdelhamid Mellouk, un mentor qui m'a accompagné tout au long de ces années dans ce projet en me laissant à la fois ce qu'il me fallait de liberté tout en m'assurant d'une disponibilité totale et d'un soutien pour toutes les actions menées et les initiatives entreprises.

Je n'oublie pas mon collègue Said Hoceini, avec qui j'ai progressé au début de ma carrière et sur lequel j'ai pu m'appuyer durant les moments de doute.

C'est donc tout naturellement, qu'une fois devenu maître de conférences, j'ai essayé de reproduire cette même démarche. Certes, cela me demandait beaucoup plus d'investissement et de temps, mais cela m'a permis de voir évoluer les sujets initiaux, émanation de mes propres réflexions, en une orientation de recherche personnelle et propre à chacun de mes étudiants. Cette démarche, qui avait pour objectif de préparer ces derniers à leurs futures carrières de chercheurs m'a également permis de m'ouvrir à d'autres champs de recherche possibles.

Je suis conscient que cette approche, aussi intéressante soit-elle, en plus de nécessiter un effort important de la part de l'encadrant, a ses limites et ne s'adapte pas à tous les profils. J'ai eu la chance de n'avoir participé qu'à l'encadrement d'étudiants de talent particulièrement réceptifs à cette méthode. Je tiens à ce sujet à remercier vivement ces jeunes chercheurs et particulièrement ceux que j'ai pu co-encadrer en thèse de doctorat : Lamine Amour, Fetia Bannour, Van Van Tong.

1. Clark, David D., et al. "A knowledge plane for the internet." Proceedings of the 2003 conference on Applications, technologies, architectures, and protocols for computer communications. 2003.

Au-delà de la méthodologie de travail, je crois que la réussite d'une telle approche tient également dans la méthode de management. Celle-ci doit se baser sur une approche coopérative et collaborative où chaque chercheur, permanent ou non-permanent, y trouve sa place. Contrairement à une méthode élitiste, cette méthode consensuelle est certes chronophage, mais permet in fine de tirer l'ensemble du groupe vers le meilleur.

La recherche universitaire, comme je la conçois et à laquelle j'aspire, doit :

- penser le monde de demain mais sans oublier les applications d'aujourd'hui,
- être reproductible, ouverte et disponible gratuitement,
- être validée par les pairs avec une méthodologie en aveugle ou en double-aveugle (la relecture des pairs pourrait/devoir être publique),
- être financée convenablement et équitablement.

L'HDR est pour moi une occasion de faire une rétrospective sur mes travaux et de confronter ma méthodologie d'encadrement à l'oeil expert de mes pairs. Cette étape marquera, sans aucun doute, une nouvelle page de ma vie de chercheur.

Pour terminer ce préambule, je tenais à vous remercier,

Vous,

mes chers collègues du laboratoire LISSI et de l'équipe TINCNet,
mes chers collègues du département RT de l'IUT de Créteil/Vitry et de l'EPISEN,
mes chers doctorants,
mes chers étudiants.

J'ai essayé de vous rendre hommage dans mon écrit avec plus au moins de succès et j'espère que chacun trouvera dans ce manuscrit une petite part de lui même.

À ma famille,
À mes magnifiques parents Abdelaziz et Faouzia,
À ma merveilleuse femme Sondes,
À mes chers enfants Sarah et Sinan.

Chapitre 1

Introduction Générale

« We are all now connected by the Internet, like neurons in a giant brain. »

Stephen Hawking

Sommaire

1.1 Brève histoire des réseaux intelligents	4
1.2 Sélection de mes activités de recherches	4
1.3 Organisation du manuscrit	6

Objectif de cette thèse d'habilitation

Cette thèse d'habilitation est rédigée sur la base de mes travaux de recherches réalisés depuis l'obtention de ma thèse de doctorat. Elle décrit mes principales contributions dans la thématique des réseaux intelligents. Elle met en avant également ma méthode et ma philosophie de travail, fruit de mon expérience de ces dernières années. Le document se termine par la présentation de certaines perspectives futures.

Je tiens à préciser que les contributions présentées dans ce manuscrit sont le fruit d'un travail collaboratif avec d'autres collègues. Une contribution dans les domaines de l'Informatique et du Génie Informatique présente de nombreuses facettes : l'étude de l'état de l'art, la proposition d'idées, la conception, la programmation, la validation, l'analyse formelle, la rédaction des articles, la relecture, la supervision ... Mon rôle n'a certainement pas été le même dans chacune d'entre elles et il est difficile de l'identifier clairement, mais tout au long de ces années, j'ai contribué à chacun de ces aspects.

L'objectif de ce manuscrit est de présenter mes activités de recherche et mes réflexions. Le document se veut, dans la mesure du possible, accessible à tout informaticien ou architecte réseaux qu'il ait une connaissance avancée ou pas de mes domaines d'intérêt. Certains détails techniques et scientifiques n'ont pas été développés dans un souci de simplification. Pour chaque contribution, l'accent est mis sur le contexte qui y a conduit, sur ses avantages ainsi que sur l'impact que cela pourrait avoir sur la thématique, mais également sur les pistes d'amélioration et les expériences à en tirer. Le cas échéant, une référence à l'une de mes communications scientifiques est donnée pour le lecteur qui souhaiterait avoir plus de détails.

1.1 Brève histoire des réseaux intelligents

L'homme a toujours cherché des moyens innovants et de plus en plus élaborés pour communiquer : les signaux de fumées, les pigeons, les courriers postaux, le télégraphe, le téléphone, la radiodiffusion et, de nos jours, Internet. Internet a révolutionné les usages, les comportements et les communications comme jamais auparavant. Cette révolution technologique, dont a rêvé "Paul Otlet" [Rayward, 2012] et qui est le résultat d'une collaboration initiée au début des années soixante [Leiner et collab., 2009], n'a toujours pas fini d'évoluer. En effet, Internet doit d'une part, s'adapter aux nouveaux services et usages et d'autre part, s'adapter au nombre croissant de dispositifs connectés qui devrait atteindre 29,3 milliards d'appareils en 2023 soit 3,6 dispositifs en réseau par habitant [Cisco, 2020]. Les politiques d'adaptation vont d'un simple surdimensionnement des infrastructures, politique qui a quelques limites [Sunny et collab., 2019], à des mécanismes de contrôle et de gestion beaucoup plus complexes basés sur des techniques d'intelligence artificielle [Robertazzi et Shi, 2020]. Appliquer les techniques d'intelligence artificielle aux réseaux de communication n'est pas un concept nouveau [Qi et collab., 2007]. Ainsi dès 1988, [Ross et collab., 1988] ont proposé un prototype de gestion de réseau reposant sur un système expert à base de règles afin de prendre des mesures de contrôle appropriées. En 1989, [Covo et collab., 1989] ont utilisé les réseaux de neurones en plus des systèmes experts pour concevoir un prototype de gestion de réseau en temps réel. À mon sens, la proposition qui a définitivement ancré le besoin d'une architecture basée sur l'intelligence artificielle est "A Knowledge Plane for the internet" [Clark et collab., 2003]. Ils ont en effet introduit la notion du "Plan de Connaissance" ou "Knowledge Plane(KP)" qui constitue un changement fondamental de la conception du réseau de communication en évoluant **d'un paradigme "centré réseau", à un paradigme "centré connaissance"**. Cette architecture offre de nombreux avantages tels que l'automatisation de la gestion, de l'optimisation et du dépannage des réseaux. Néanmoins, peu de prototypes mettant en œuvre cette proposition ont été déployés. La cause principale à cela fut sans doute le manque de programmabilité des équipements. **L'émergence de la virtualisation des réseaux et du paradigme SDN (Software-Defined Networking)** qui découplent le plan de contrôle du plan de données a levé ce verrou [Feamster et collab., 2014]. En effet, le SDN, par une centralisation du contrôle basée sur une vue globale et unique du réseau, permet de programmer dynamiquement et facilement les différents éléments pour répondre à un certain contexte. Mes activités de recherche s'inscrivent dans ce cadre.

1.2 Sélection de mes activités de recherches

Mes activités de recherche portent principalement sur le paradigme des réseaux intelligents et plus particulièrement sur les mécanismes de contrôle, de supervision et d'automatisation dans un contexte de "très grande échelle". Pour mieux comprendre mon domaine d'intérêt, je propose un bref éclairage sur chacun de ces termes.

Les Réseaux Informatiques

Un réseau informatique est un ensemble de composants interconnectés et qui permettent la communication entre des ordinateurs ou tout autre objet numérique connecté (par exemple, smartphones, capteurs, drones, ...). Ces réseaux peuvent être filaires ou sans fil et sont de taille et de portée différentes : allant des BAN (Body Area Networks)

et du PAN (Personal Area Networks) au WAN (Wide Area Networks) en passant par le LAN (Local Area Networks).

Dans notre monde qui est de plus en plus connecté, les réseaux informatiques n'ont cessé d'évoluer afin de répondre à un besoin de plus en plus croissant à la fois en termes de besoin et d'usage. Les dernières évolutions que représente les paradigmes SDN (Software-Defined Networking) et la NFV (Network Function Virtualization) ont apporté la flexibilité et ont amélioré l'efficacité et la maintenance en permettant de dissocier le contrôle du matériel (qui étaient jusque là spécialisé et propriétaire). Mes travaux se basent essentiellement dans ce contexte.

Contrôle des réseaux à large échelle

La centralisation de la logique de contrôle dans les réseaux SDN est un frein au passage à l'échelle. En fait, lorsque la taille du réseau augmente (par exemple, commutateurs, hôtes, etc.), le contrôleur SDN centralisé est fortement sollicité et donc surchargé (en termes de bande passante, de puissance de traitement et de mémoire). En outre, lorsque la taille du réseau augmente, les délais de communication entre le contrôleur SDN et les commutateurs de réseau peuvent être allongés, ce qui affecte les latences des flux. Une architecture basée uniquement sur un seul contrôleur SDN physiquement centralisé peut ne pas répondre aux exigences de performance des réseaux à grande échelle. Il est donc important de concevoir des architectures, certes logiquement centralisées pour garder les apports du SDN, mais qui soient physiquement distribuées.

La supervision (Monitoring)

Au début des réseaux informatiques et d'Internet, la supervision des réseaux était une thématique inconnue. Si l'on rencontrait un problème de réseau, il suffisait de lancer quelques paquets de données pour localiser la source du problème et y remédier. Le 27 octobre 1980, un événement inhabituel s'est produit et a provoqué le dysfonctionnement du réseau ARPANET pendant plusieurs heures ce qui a donné lieu aux premières discussions sur l'intérêt de la supervision et de la mise en place d'outils de gestion du réseau. La supervision du réseau est l'utilisation d'un système qui surveille en permanence un réseau informatique pour détecter des comportements anormaux et qui informent l'administrateur ou déclenche des correctifs en cas de panne. Au fur et à mesure de l'évolution des réseaux, la supervision est devenue de plus en plus élaborée et autonome.

L'automatisation

Il s'agit de concevoir des méthodes et des approches qui permettent d'automatiser le fonctionnement du réseau de manière à coordonner les composants matériels et logiciels nécessaires à l'exécution d'une application ou d'un service. Un objectif important est d'automatiser l'exécution des requêtes réseau et de réduire la quantité d'intervention humaine nécessaire pour fournir une application ou un service.

Quelques mots sur les réseaux intelligents

Principales contributions

Cette thèse met en évidence une **sélection de mes contributions scientifiques**. Cette sélection comprend les trois sujets suivants : **l'architecture et la faisabilité des réseaux**

Le chapitre 6 conclut cette thèse. Toutes mes publications coécrites sont listées dans le chapitre 7.

Bibliographie

- Cisco. 2020, «Cisco annual internet report (2018–2023), white paper.», URL "<https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html>". 4
- Clark, D. D., C. Partridge, J. C. Ramming et J. T. Wroclawski. 2003, «A knowledge plane for the internet», dans *Proceedings of the 2003 conference on Applications, technologies, architectures, and protocols for computer communications*, p. 3–10. 4
- Covo, A. A., T. M. Moruzzi et E. D. Peterson. 1989, «Ai-assisted telecommunications network management», dans *1989 IEEE Global Telecommunications Conference and Exhibition 'Communications Technology for the 1990s and Beyond'*, p. 487–491 vol.1. 4
- Feamster, N., J. Rexford et E. Zegura. 2014, «The road to sdn : An intellectual history of programmable networks», *SIGCOMM Comput. Commun. Rev.*, vol. 44, n° 2, doi :10.1145/2602204.2602219, p. 87–98, ISSN 0146-4833. URL <https://doi.org/10.1145/2602204.2602219>. 4
- Leiner, B. M., V. G. Cerf, D. D. Clark, R. E. Kahn, L. Kleinrock, D. C. Lynch, J. Postel, L. G. Roberts et S. Wolff. 2009, «A brief history of the internet», *ACM SIGCOMM Computer Communication Review*, vol. 39, n° 5, p. 22–31. 4
- Qi, J., F. Wu, L. Li et H. Shu. 2007, «Artificial intelligence applications in the telecommunications industry», *Expert Systems*, vol. 24, n° 4, p. 271–291. 4
- Rayward, W. B. 2012, «Paul otlet, an encounter», *Cahiers de la documentation–Bladen voor documentatie*, vol. 2, p. 71. 4
- Robertazzi, T. G. et L. Shi. 2020, «Machine learning in networking», dans *Networking and Computation*, Springer, p. 151–190. 4
- Ross, M., A. Covo et C. Hart. 1988, «An ai-based network management system», dans *IEEE International Performance Computing and Communications Conference*, IEEE Computer Society, Los Alamitos, CA, USA, p. 458,459,460,461, doi :10.1109/PCCC.1988.10121. URL <https://doi.ieeecomputersociety.org/10.1109/PCCC.1988.10121>. 4
- Sunny, A., R. El-Azouzi, A. Arfaoui, E. Altman, S. Poojary, D. Tsilimantos et S. Valentin. 2019, «Enforcing bitrate-stability for adaptive streaming traffic in cellular networks», *IEEE Transactions on Network and Service Management*, vol. 16, n° 4, p. 1812–1825. 4

Chapitre 2

Curriculum Vitae

« Mieux vaut être un diamant avec un défaut qu’être un caillou parfait »

Confucius

Resumé

De 2010 à 2013, j’ai préparé mon doctorat en Informatique et Génie Informatique à l’école doctorale MSTIC de l’Université Paris Est. Après l’obtention de mon doctorat, j’ai été recruté pendant une année (2013-2014) en tant qu’attaché temporaire d’enseignement et de recherche. Motivé à la fois par l’enseignement et par la recherche, je me suis présenté en 2014, à l’âge de 28 ans, au concours de recrutement de Maître de conférences et j’ai été lauréat d’un poste (section 61) ouvert à l’Université Paris Est Créteil. Mes travaux de recherche portent sur les mécanismes adaptatifs dans les réseaux dynamiques à grande échelle. Ces mécanismes sont basés sur des connaissances contextualisées, sur la virtualisation des fonctions réseau (Network Functions Virtualization : NFV) et sur les réseaux pilotés par logiciel (Software-Defined Networking : SDN). Durant ces 8 dernières années, j’ai co-encadré 3 thèses de doctorat soutenues ainsi que plusieurs étudiants de niveau M2. J’ai également participé activement à l’effort global au sein de mon équipe dans le cadre de projets collaboratifs de recherche. **Ces actions me permettent aujourd’hui d’envisager le dépôt de l’habilitation à diriger des recherches (HDR)**

Sommaire

2.1 Présentation synthétique des thématiques de recherche	10
2.2 Encadrement doctoral et scientifique	10
2.3 Diffusion, rayonnement et vulgarisation	12
2.4 Responsabilités scientifiques	18
2.5 Responsabilités au niveau local	20
2.6 Participation aux conseils des composantes	20

2.1 Présentation synthétique des thématiques de recherche

Les travaux que je mène s’inscrivent dans le cadre du paradigme dit des Next Generation Networking and Internet (NGNI) et particulièrement autour de deux problématiques :

- Comment intégrer le ressenti de l’usager dans le contrôle du réseau?
- Comment faire face à la forte dynamique de certains réseaux ainsi qu’à la très grande échelle?

Afin de rendre les approches proposées déployables à large échelle, j’ai travaillé, dans le cadre d’une thèse, que j’ai co-encadré (soutenue au mois de novembre 2019), sur la mise en oeuvre d’un plan de connaissances distribué bâti sur le paradigme des réseaux pilotés par logiciel (Software Defined Networking : SDN). Ce plan a pour finalité de disséminer l’ensemble des connaissances contextuelles du réseau afin de permettre le développement de nouveaux mécanismes efficaces de contrôle du réseau (ces mécanismes traitent principalement des fonctions de routage et de gestion dynamique de ressources). La nature des connaissances contextuelles prises en charge par ce plan intègre celle relative à la Qualité de l’Expérience (QdE ou QoE : Quality of Experience). Cette mesure, qui ne se focalise pas sur la qualité intrinsèque du réseau, mais plutôt sur la qualité telle qu’elle est perçue par l’usager, a fait l’objet d’une seconde thèse de doctorat soutenue au mois de décembre 2018. Cette thèse, que j’ai co-encadré, a eu pour sujet l’étude des méthodes qui permettent d’estimer la métrique de QdE dans le contexte des applications mobiles. Dans la continuité de ces 2 thèses, j’ai co-encadré une 3e thèse, soutenue en décembre 2021, ayant pour objectif de proposer de nouveaux mécanismes de dépannage pour les réseaux d’opérateurs. Enfin, sur un tout autre plan, en lien avec le positionnement scientifique de mon laboratoire, accès en grande partie sur des technologies pour la santé, je travaille sur des dispositifs et approches intelligents qui tentent de lever certains verrous technologiques inhérents à ce domaine ; par exemple, j’ai été amené à réfléchir sur la mise en oeuvre de dispositifs multi-interfaces qui, sur la base d’informations contextuelles, choisissent automatiquement et à la volée la meilleure interface de sorties pour chaque service de télésanté assurant ainsi la meilleure QdE possible. J’ai également initié une réflexion sur l’apport des applications mobiles dans la lutte contre le Covid-19.

2.2 Encadrement doctoral et scientifique

Thèses soutenues

M. Lamine AMOUR, Mise en œuvre d’un modèle adaptatif pour l’estimation de la qualité réelle perçue par l’usager (QoE) : Application aux services mobiles

- Date de soutenance : 3 décembre 2018.
- Financement : Contrats FUI POQEMON. Taux : 50%. sous la direction du Professeur Abdelhamid Mellouk (UPEC).
- Jury de thèse (rapporteurs soulignés) : Stéphane Maag, Adlen Ksentini, Thierry Divoux, Ana Rosa Cavali, Samia Bouzefrane, Sami Souihî, Abdelhamid Mellouk.
- Production des travaux de la thèse : 1 article de rang A [[Amour et collab., 2020](#)]+ 7 articles de rang B [[Amour et collab., 2015c, 2016, 2015a,b, 2017, 2018, 2019](#)]
- Situation actuelle : Enseignant-chercheur recruté en CDI à ESME-SUDRIA

Mme Fetia BANNOUR : Contributions pour le contrôle distribué dans les réseaux SDN.

- Date de soutenance : 19 novembre 2019.
- Financement : Projets FUI POQEMON, Celtic+ NOTTS et contrat ATER en 2019. Taux : 50%, sous la direction du Professeur Abdelhamid Mellouk (UPEC).
- Jury (rapporteurs soulignés) : Olivier Festor, Gérardo Rubino, Samia Bouzefrane, Thierry Divoux, Djamel Zeghlache, Fabrice Guillemin, Sami Souihi, Abdelhamid Mellouk.
- Production des travaux de la thèse : 2 articles de rang A [[Bannour et collab., 2017a, 2020](#)] + 3 articles de rang B [[Bannour et collab., 2017b, 2018, 2019](#)]
- Situation actuelle : Maître de Conférences nommée en septembre 2020 à l'ENSIIE, Univ. Evry de l'Université Paris Saclay.

M. Van Van TONG, Adaptive and scalable model for preventing and troubleshooting network issues in the context of encrypted data traffic

- Date de soutenance : 13 décembre 2021.
- Financement : Bourse d'excellence obtenue dans le cadre du programme Franco-Vietnamien (2018-2021). Taux : 50%, sous la direction du Professeur Abdelhamid Mellouk (UPEC)..
- Jury (rapporteurs soulignés) : Jean-philippe Georges, Farid Naït-Abdesselam, Houda Labiod, Abderrahim Benslimane , Djamel Zaghache , Sami Souihi, Abdelhamid Mellouk
- Production des travaux de la thèse : 1 article de rang A [[Tong et collab., 2021b](#)] et 7 articles de rang B [[Van et collab., 2018a](#); [Tong et collab., 2018](#); [Van et collab., 2018b](#); [Amour et collab., 2019](#); [Chu et collab., 2019](#); [Van et collab., 2020](#); [Tong et collab., 2021a](#)].
- Situation actuelle : Maître de Conférences nommé en septembre 2022 à l'HUST, Univ. polytechniques de Hanoi, Vietnam.

Encadrement de Stages Doctoraux

Mlle Jihen Bennaceur (3 mois : 03/04/2017 à 30/06/2017)

- Sujet : Apport des réseaux SDN dans la radio cognitive et collecte de donnée sécurisée.
- Production scientifique : 1 article de Rang B [[Bennaceur et collab., 2017](#)]

Encadrement de Stages de Master 2 et d'Ingénieurs

M. Abdelraouf Ghamri (02/2016 à 07/2016)

- Sujet : Les mécanismes de sélection d'interfaces dépendant du contexte.
- Taux : 50%

M. Skander Mhadhbi (03/2016 à 09/2016)

- Sujet : Les mécanismes de compression/décompression d'objets 3D ainsi que sur les mécanismes adaptatifs de streaming d'objets 3D sensible à la QoE.
- Taux : 50%

M. Raouf Haouili (5 mois : 01/02/2017 à 30/06/2017)

- Sujet : L'interopérabilité au sein d'un même cluster de 2 contrôleurs SDN différents (ONOS et OpenDayLight)
- Taux : 50%

M. Tuan Chu (5 mois : 01/02/2019 à 28/06/2019)

- Sujet : Mise en oeuvre d'une plateforme d'expérimentation bâtie sur le paradigme des réseaux pilotés par le logiciel (SDN). Ce stage a fait l'objet d'une publication dans un congrès publié.
- Taux : 50%
- Production scientifique : 1 article de Rang B [Chu et collab., 2019]

M. Hai Nam Nguyen (6 mois : 01/06/2020 à 30/11/2020)

- Sujet : Mise en oeuvre d'une architecture SDN hybride et sécurisée
- Taux : 50%
- Production scientifique : 2 article de Rang A [Nam Nguyen et collab., 2021; Hoang et collab., 2022] et un article de Rang B [Nguyen et collab., 2022a]

M. Nicolas Kalinkine (4 mois : 01/09/2020 à 23/12/2020)

- Sujet : Mise en oeuvre d'une architecture distribuée et scalable de gestion des connaissances.
- Taux : 50%
- Production scientifique : 1 Plateforme en cours de validation

Tharsana Balachandran (6 mois : 01/02/2021 à 30/08/2021)

- Sujet : Prediction of Escherishia coli concentration with transfer learning
- Taux : 30%
- Production scientifique : 1 article de Rang B [Balachandran et collab., 2022]

M. Do. Quang Huy (6 mois : 01/09/2021 à 30/04/2022)

- Sujet : A Scalable Blockchain for the SDN
- Taux : 50%
- Production scientifique : 1 article de Rang B [Do et collab., 2022]

M. Tien Anh Nguyen (6 mois : 01/09/2021 à 30/04/2022)

- Sujet : NLP and Blockchain for Networking
- Taux : 30%
- Production scientifique : 1 article de Rang B [Nguyen et collab., 2022b]

2.3 Diffusion, rayonnement et vulgarisation

Organisation (Co) de Conférences Internationales

- Panel co-Chair (Sami Souihi, Minh Jo et Juan Corchado) IEEE HPSR 2021
 - Rôle : Organisation des sessions « panel », choix des thématiques, sélection des intervenants pertinents
 - Lieu (Organisation Hybride) : Paris, France et virtuelle
 - Date : 07-09 Juin 2021

- TPC co-Chair (Sami Souihi, Salim Bitam et Scott Fowler) 7th IEEE Smart Communications in Network Technologies 2018 (IEEE SACONET'18)
 - Rôle : Organisation des relectures, Échanges avec les auteurs, Organisation du programme, Organisation des sessions, Animation de sessions
 - Lieu : EL Oued, Algerie
 - Date : 29-31 octobre 2018
 - Nombre de sessions : 11
 - Nombre de papiers soumis 150
 - Nombre de papiers acceptés : 59
 - Nombre de participants : 155
- General co-Chair (Sami Souihi et Hanen Idoudi) 2018 IEEE 29th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (IEEE-PIMRC) - Workshop WS-08 "International workshop of Trends and challenges in Social Internet of Things (TC-SIoT'2018)
 - Rôle : Rédaction et soumission de la proposition du Workshop, Publicité, Organisation des relectures, Échanges avec les auteurs, Organisation de la session, Animation de la session
 - Lieu : Bologne, Italie
 - Date : 9 septembre 2018
 - Nombre de sessions : 1 (durée : 1h30)
 - Nombre de papiers soumis 10
 - Nombre de papiers acceptés : 3 + une keynote
 - Nombre de participants : 20
- Symposium TPC co-Chair (Sami Souihi, Kohei Shiimoto et Dzimitry Kliazovich) - Communications QoS, Reliability and Modeling Symposium(CQRM) -IEEE International Conference on Communications 2017 (IEEE ICC) (IEEE Flagship Communications Society Conference).
 - Rôle : Organisation des relectures, Échanges avec les auteurs, Organisations des sessions
 - Lieu : Paris, France
 - Date : 21-25 Mai 2017
 - Nombre de sessions : 9
 - Nombre de papiers soumis 131
 - Nombre de papiers acceptés : 48

Chair de sessions de Conférences Internationales

- Chairman de la session «Quality Of Service», IEEE Global Communications Conference (GLOBECOM 2022), Rio de Janeiro, Bresil, sous la présidence de Lotfi Mhamdi, Qi Li et Selma Boumerdassi (TPC-co chairs du symposium NEXT-GENERATION NETWORKING AND INTERNET).
 - Durée : 1h30
 - Nombre de papiers présentés : 5

- Chairman de la session «Network Virtualization», IEEE Global Communications Conference (GLOBECOM 2021), Madrid, Espagne, sous la présidence de Hassine Moun- gla, Michele Nogueira et Shaoen Wu (TPC-co chairs du symposium Communica- tion Software and Multimedia).
 - Durée : 1h30
 - Nombre de papiers présentés : 5
- Chairman de la session «MC2 (Software and Virtual Networks 1) », International Conference on Network and Service Management (CNSM 2020), Virtual, 2020 sous la présidence de Muge Sayit et Stuart Clayman (TPC-co chairs).
 - Durée : 1h
 - Nombre de papiers présentés : 4
- Chairman de la session «SAC IoT5 (Localization and Tracking I) », IEEE Global Com- munications Conference (GLOBECOM 2019), Waikoloa, HI, USA, 2019 sous la pré- sidence de Ridha Soua et Pietro Manzoni (TPC-co chairs du SAC – IoT).
 - Durée : 1h
 - Nombre de papiers présentés : 5

Organisation de colloques scientifiques

- Co-président (Sami Souihi et Hanen Idoudi) de l'école d'automne Blockchain'19.
 - Rôle : Rédaction et proposition de l'évènement, Publicité, Échanges avec les intervenants, Organisation de l'évènement, Animation de sessions
 - Lieu : Université de la Manouba, Tunisie
 - Date : 27 au 29 Novembre 2019
 - 120 participants, 6 keynotes, 5 ateliers pratiques et une session scientifique.
- Co-organisation du colloque national SDN Day 2018, porté par les GDR CNRS MACS (ARC) et ASR (ResCom), ayant pour thème les « IDNs-Intelligence Defined Net- works» (Co-organisé par l'équipe CIR/LISSI, Huawei, INRIA, Orange Labs et IRT SystemX)
 - Rôle : (co) Organisation de l'appel à contributions, publicité, invitation d'in- tervenants, (co) organisation et (co) animation de l'évènement.
 - Lieu : Paris, France
 - Date : 22 novembre 2018
 - 124 participants, 8 keynotes, 14 posters.

Membre des comités scientifiques des revues internationales suivantes

- Membre du Comité thématique de la revue MDPI Electronics (ISSN 2079-9292; CO- DEN : ELECGJ) depuis décembre 2020. Ma tâche principale est de diriger, périodi- quement, les numéros spéciaux liés à mes domaines de recherche et de suggérer de nouvelles thématiques et d'inviter des chercheurs en tant que rédacteur invité pour ces nouveaux sujets. À ce titre, j'ai proposé une édition spéciale à paraître en 2021 (cf. section 4.11).

Membre des comités scientifiques des conférences internationales suivantes (5 dernières années)

- IEEE International Conference on Communications (IEEE ICC), NGNI 2023, Rome, Italie, 28 Mai - 01 juin 2023. (Rapporteur de 2 papiers).
- IEEE Global communications conference (IEEE GlobeCom, IEEE Flagship Communications Society Conference), Communications Software, Signal Processing for Communications Symposium 2022, Rio de Janeiro, Bresil, 04-08 December 2022. (Rapporteur de 5 papiers).
- International Conference on High-Performance Switching and Routing (HPSR 2021), Paris, France, 07-10 Juin 2021. (Rapporteur de 6 papiers).
- IEEE Global communications conference (IEEE GlobeCom, IEEE Flagship Communications Society Conference), Communications Software, Signal Processing for Communications Symposium 2019, Waikoloa, HI, USA, 9-13 December 2019. (Rapporteur de 6 papiers).
- IEEE Global communications conference (IEEE GlobeCom, IEEE Flagship Communications Society Conference), Communications Software, Next-Generation Networking & Internet Symposium 2018, Abu Dhabi, UAE, 9-13 December 2018. (Rapporteur de 6 papiers).
- IFIP International Conference on Performance Evaluation and Modeling in Wired and Wireless Networks (PEMWN) 2018, Toulouse, France, 26-28 September 2018. (Rapporteur d'un papier).
- IFIP International Conference on Performance Evaluation and Modeling in Wired and Wireless Networks (PEMWN) 2017, Paris, France, 28-30 novembre 2017. (Rapporteur de 3 papiers).
- IEEE International Conference on Communications (IEEE ICC), CQRM & NGNI 2017, Paris, France, 21-25 Mai 2017. (Rapporteur de 6 papiers).

Rapporteur (régulier) pour les revues scientifiques suivantes

- IEEE Communications Letters, depuis 2018 (3 papiers/an)
- Future Generation Computer Systems (science direct), depuis 2020 (2 papiers/an)
- MDPI Sensors, depuis 2018 (1 papier/an)
- International Journal of Communication Systems (Wiley & Sons), depuis 2014 (3 papiers/an)
- Rapporteur irrégulier pour les revues (Journal on Selected Area on Communication (JSAC), IEEE Transactions on Networking, IEEE Transactions on Communications)

Membre élu au chapitre Français de l'IEEE Communications Society (ComSoc)

- 2014 à ce jour
- Création et pilotage en 2016 d'une récompense pour les 3 meilleurs travaux de recherche menés par des étudiants dans le cadre de leur stage de fin d'études (Stage de Master 2 ou d'école d'ingénieurs) effectué en France (ou à l'étranger si l'étudiant est inscrit dans un établissement français).

Séjours dans des laboratoires/universités étrangers

- 2019 : École Nationale des Sciences de l'Informatique de Tunisie. Collaboration avec H. Idoudi, Maitresse de conférences,
 - Objectif : Organisation de l'école d'automne Blockchain'19 et l'enseignement d'un module de Cloud Networking pour les étudiants du Master 2.
 - Durée du séjour : 2 semaines
- 2018 : School of Information and Communication Technology, Hanoi University of Science and Technology, Vietnam. Collaboration avec H. A. Tran, Associate Professor, dans le cadre du projet Sense-Sud.
 - Objectif : Collaboration et publication d'un article de rang A (cf. [Revue 3])
 - Durée du séjour : 10 jours

Participation jurys de thèse (hors établissement)

Soutenance de thèse

- Doctorant : M. Loïc Desgeorges
Directeur de Thèse : Thierry Divoux et de Jean-Philippe Georges
Président du jury : Didier Maquin
Rapporteurs : Isabel Demongodin, Damien Magoni, Pascal Berthou.
Examineurs : Pascal Berthou, Sami Souihi, Ye-Qiong Song.
Date de soutenance : 17-11-2022
Sujet : Détection d'anomalies de sûreté et sécurité d'un contrôle centralisé de réseau
Rôle : Examineur.

Soutenance de mi-parcours

- Doctorant : M. Thierno Sambe
Directeur de Thèse : Stéphane Maag
Co-encadrants : Ana Cavalli, Hamet Seydi et Karim Konate
Date de première inscription : 20/12/2017
Date de l'évaluation mi-parcours : 30/08/2019
Sujet : Monitoring Avancé et Hétérogène des ERP pour les processus de production industriels.
Rôle : Membres du Comité d'évaluation (avec Pr. Jean-Marc Delosme)

Actions de vulgarisation scientifique & Keynotes

- Naji Bouchiba, Nelly Chatue Diop, Sami Souihi, Hichem Bourak, Matthieu Quiniou, Xhumanisa : Des solutions de santé numériques pour l'Afrique, Smart Summit Ocean Indien, 22-24 mai 2020, Invitation par Marc Agouni (comité d'organisation).
- Tutoriel présenté dans le cadre de l'école d'automne SaCoNeT 2019, Oran, Algérie, 23 décembre 2019 (180 participants) : JAIN : Just Another Intelligent Network, invité par Abdelmadjid Snouber et Abdelhamid Mellouk (comité d'organisation).

- Sami Souihi, La Blockchain : démystification histoire et perspectives, Les "Mercredis du numérique", La Plaine Commune, Grand Paris, France. mercredi 3 avril 2019. Les "Mercredis du numérique" ont pour ambition de partager expériences, projets, pratiques et de susciter réflexions et échanges entre élus, agents de Plaine Commune et intervenants afin de mieux comprendre les enjeux et perspectives concernant le numérique. Invité par Marc Bourgarel (comité d'organisation).
- Tutoriel présenté dans le cadre de l'école d'automne autour du développement numérique par les communications intelligentes, El Oued, Algerie, 27 octobre 2018 (220 participants) : les réseaux logiciels (Software-defined networking), une révolution?, invité par Abdelhamid Mellouk (comité d'organisation).
- A. Mellouk, S. Souihi, A software defined Cognitive Radio Networks : An application to IoMT, IoT/WSN workshop, USTH, Hanoi, Vietnam, 21-22 Feb, 2017. Invitation par les organisateurs du Workshop dans le cadre de l'université Franco-Vietnamienne (USTH).
- A. Mellouk, S. Souihi, State-Dependent Time Evolving Multi-Constraint Control : Resources allocation, Grid and High Performance Computing workshop, USTH, Hanoi, Vietnam, 23-24 Feb, 2017 Invitation par les organisateurs du Workshop dans le cadre de l'université Franco-Vietnamienne (USTH).
- Organisation des 1ere, 2e et 3e Édition EduPi(100 participants par évènement) 23 septembre 2017, 22 octobre 2018, 30 novembre 2019.

Edupi est une action éducative initiée par moi-même et co-organisée par les membres de l'équipe CIR/LISSI (TincNET) ainsi que d'autres partenaires (CEA LIST, Chaire ITEN/Unesco, Université de la Manouba, Université de Jendouba, ...). Elle est concrétisée par un évènement annuel qui se tient sous la forme d'un séminaire d'échanges entre étudiants, professeurs et industriels. Cette action a pour ambition d'aider à la réduction des inégalités technologiques en proposant un séminaire scientifique et thématique de qualité dans des zones reculées. La journée est organisée en 2 parties, une première partie qui se veut vulgarisée et une 2e partie sous forme de table ronde. Ayant pour thème les objets connectés, la première édition d'Edupi s'est tenue le 23 septembre 2017 à l'Institut Supérieur des Études Technologiques du Kef en Tunisie. La seconde édition a eu lieu quant à elle à l'université de Jendouba en Tunisie. Cette édition 2018 s'est concentrée sur les thématiques autour de «La santé connectée». La 3e édition (2019) s'est intéressée au thème de « l'agriculture numérique et connectée ». Il n'y a pas eu d'éditions de 2020 à 2022 à cause du Covid-19 et du contexte post-déconfinement.

- Co-organisation d'un Webinaire « Towards eHealth in post-containment era : epidemics experiences from Africa and Europe »(40 participants), 09 juillet 2020.
 - Durée : 2h30.
 - 3 keynotes et une table ronde.

Activités éditoriales

- Guest Editor (avec Salim Bitam, Abdelhamid Mellouk, Thiago Abreu, Said Hoceini, Scott Fowler, Saci Medileh, Swades De & Abdallah Shami) de la special Issue "Smart Communications for Autonomous Systems in Network Technologies" de la revue IJCS WILEY
 - Impact Factor de la revue : 1.099

- Guest Editor (avec Scott Fowler & Hai Anh Tran) de la special Issue "Advances in Communications Software and Services" de la revue MDPI Electronics (ISSN 2079-9292; CODEN : ELECGJ)
 - Impact Factor de la revue : 2.412

2.4 Responsabilités scientifiques

Contrats de recherche co-portés financés

- PoQEMoN (Platform for Quality Evaluation of Mobile Networks) : Projet FUI16, Pôle de compétitivité Systematic, Groupe Thématique Telecom : 175 k€

Date : 01/2014-06/2016

Membres du consortium : Ip-label, Genymobile, Data Publica, Next Interactive Media (Groupe Next Radio), Orange, Altran, Université Paris Est-Créteil (UPEC), INRIA, HEC. Porteur du projet : Abdelhamid Mellouk.

Responsable du suivi technique : Sami Souihi

PoQEMoN a pour ambition de proposer un système de référence dans l'évaluation de la qualité des réseaux mobiles. Par qualité, nous entendons : couverture, disponibilité et performance des réseaux sur des usages représentatifs (data, vidéo). En l'absence de normes de mesure de la qualité des réseaux et des services mobiles, PoQEMoN a pour objectif de définir un référentiel permettant l'évaluation des réseaux et leur comparaison (benchmark). Ce référentiel peut être mis en œuvre quelle que soit la technologie mobile utilisée (2G, 3G, 3G+, 4G, 5G, etc..) et peut donc être transposable dans l'ensemble des pays de l'Union Européenne. En tant que responsable technique chargé du suivi du projet, et sous l'autorité du Professeur Abdelhamid Mellouk, j'ai coordonné pour ma part les travaux liés à la mise en œuvre et au suivi des tâches qui nous sont affectées au sein du consortium. Du point de vue de mes thématiques liées à la recherche, ces travaux consistent à modéliser la qualité perçue par l'utilisateur (QdE : Qualité de l'Expérience) et d'en étudier la corrélation avec les paramètres de Qualité de Service issus du réseau.

- NOTTS (Next generation over-the-top multimedia Services Technology) : projet Européen Eureka Celtic+ : 200 k€.

Date : 01/2014-04/2016

Membres du consortium : Indra Sistemas, S.A. (IND), Alcatel Lucent (ALU), Dyccec (DYC), ADTEL Sistemas de Telecomunicacion S.L. (ADTEL), Acreo (ACR), Ericsson (EABS), Alkit Communication (ALK), Procera Networks (PRO), Lund University (LTH), Institut Telecom SudParis (ITP), University Paris-Est Creteil (UPEC), IP-Label (ILN), Thomson Video Networks (TVN), VTT Technical Research Centre of Finland (VTT), Anvia Oyj (ANV), Hibox Systems Oy Ab (HIB), Networker QoS Oy (NET), Oy Omnitele Ab (OMN), Finland, Orange Polska S.A. (OPL), Portugal Telecom Inovação Sistemas (PTINS)

Porteur du projet : Abdelhamid Mellouk.

Responsables du suivi technique : Sajid Mushtaq, Brice Augustin et Sami Souihi.

Responsables de la plateforme de démonstration : Sami Souihi.

Le projet européen NOTTS concerne la gestion de la qualité de l'expérience (QoE) des services multimédias diffusés sur l'Internet, appelés services « Over The Top (OTT) », dans les réseaux opérateurs et cloud. Les approches proposées dans le

cadre de ce projet concernent l'intégralité de la chaîne de distribution des données multimédia : de l'encodage aux architectures de réseaux de distribution en passant par la supervision. L'objectif final du projet est de fournir une solution technologique, sous la forme d'un prototype, qui apporte une amélioration pour l'ensemble des parties prenantes (client final, fournisseur de service et opérateur). Dans le cadre de ce projet, j'ai assuré la co-responsabilité du suivi technique et la responsabilité du montage de la plateforme de démonstration, coté UPEC, qui a été livrée à la fin du projet en Mars 2016.

- Sense-Sud : Innovative Sensors and IoT Telecommunication Networks for Environmental Surveillance in Southern Countries : Projet financé dans le cadre de l'Institut de Recherche pour le Développement (IRD) ;
Période : 2018-2021 ;
26 partenaires issus de 4 pays (France, Cameroun, Vietnam et Sénégal) ;
Budget : 60K€,
Rôle : co-porteur coté UPEC
- HUBOmics : Plateforme distribuée de stockage de données (Santé & Environnement) et extraction automatique de méta-données : projet financé dans le cadre de l'AAP : Bonus Qualité Enseignement-Recherche (BQ-ER) de l'UPEC.
Durée : 12 mois (2020) ;
Budget : 6940 €.
Membres du consortium : 4 partenaires français : LEESU, LISSI, Gly-CRRET, IMRB
Rôle : co-porteur
- Urbapoll : Artificial intelligence and the internet of Thing to monitor urban pollution in smart cities. Projet financé dans le cadre du programme OPUR (Observatoire des Polluants Urbains).
Durée : 12 mois (2021) ;
Budget : 6652 €.
Membres du consortium : 3 partenaires français : LEESU, LISSI, LISA
Rôle : co-porteur
- HiFlow, Health Impact of Pollution on emergent patients' flow, dans le cadre de AAP "Amorçage" 2021 de Paris-Est Sup. Notre rôle porte sur le déploiement des capteurs ;
Durée de 36 mois,
Budget labo 45k€,
Membres du consortium : UPEC, Université Gustave Eiffel (Lab'Urba), ARS-IDF (Dir-Nov), Conseil Départemental 94, AP-HP (SAMU94, SAU Henri Mondor, DIM) ;
Porteurs : Matthieu Heidet, Médecin urgentiste chercheur associé au sein de CIR et Sami Souihi (MdC UPEC).
- Projet OMEGA, Optimisation systématique de la permanence des soins Ambulatoires, dans le cadre de l'AAP APRES phase2 de l'APHP, avec notre rôle portant l'étude des modèles non stationnaires.
Durée de 23 mois,
Budget labo 50€,
Membres du consortium : UPEC, AP-HP (SAMU94, SAU Henri Mondor, DIM), PDSA (Médecins à domicile, SOS médecins), EHESP ;
Porteurs : A. Mellouk (Pr UPEC), Matthieu Heidet (Médecin urgentiste) et Sami Souihi (MdC UPEC).

- Projet FORBATH : Forecasting the microbial water quality in urban inland and coastal bathing waters. Projet Eurostars (E115737).

Période : 2022-2025

6 partenaires de 3 pays (France, Portugal, Allemagne)

Budget : 1.975.059,49 € (dont 241,164 € pour le laboratoire LISSI).

Rôle : co-porteur coté UPEC.

En conclusion de cette partie, je mentionne par ailleurs que je participe très activement à l'effort global au sein de mon équipe dans le cadre du montage de nouveaux projets collaboratifs de recherche qui sont aujourd'hui en cours de dépôt.

2.5 Responsabilités au niveau local

- J'ai été responsable pédagogique du parcours ASReS (Administration et Sécurité des Infrastructures Réseau et Systèmes pour la E-Santé) de la Licence Professionnelle MERIT (métiers des réseaux informatiques et télécommunications). Ce parcours porte sur les métiers du réseau informatique et de la TechSanté. L'effectif tourne en moyenne autour de 12 étudiants/an. Période : 2015 à 2021.
- Participation très active au montage de la filière Ingénierie et la Technologie en Santé (ITS) dans le cadre de la création de la composante d'une école d'ingénieurs (EPISEN) ouverte en septembre 2016 (accord obtenu de la CTI en janvier 2016). Outre, le pilotage de certaines unités d'enseignement et le recrutement de vacataires et de conférenciers, j'ai par exemple contribué à la rédaction du dossier de demande d'accréditation de la formation qui a été transmis à la commission des titres d'ingénieur en 2016 et au dossier de renouvellement d'accréditation (2019). Je suis également, depuis 2017, référent des étudiants des 3e année et à ce titre je suis en charge du suivi du groupe, du suivi des absences, du conseil et du suivi des étudiants dans leurs différentes démarches ainsi que de la remontée des informations à l'équipe pédagogique. J'ai également co-porté la création de la filière Ingénierie et la Technologie en Santé (ITS) par apprentissage que je pilote depuis septembre 2022.

2.6 Participation aux conseils des composantes

- Membre élu du Conseil de Laboratoire LISSI (2017 à ce jour), collègue 2 (McF non HDR).

Bibliographie

Amour, L., M. I. Boulabiar, S. Souihi et A. Mellouk. 2018, «An improved qoe estimation method based on qos and affective computing», dans *2018 International Symposium on Programming and Systems (ISPS)*, IEEE, p. 1–6. [10](#)

Amour, L., M. S. Mushtaq, S. Souihi et A. Mellouk. 2017, «Qoe-based framework to optimize user perceived video quality», dans *2017 IEEE 42nd Conference on Local Computer Networks (LCN)*, IEEE, p. 599–602. [10](#)

- Amour, L., S. Sami, S. Hoceini et A. Mellouk. 2015a, «Building a large dataset for model-based qoe prediction in the mobile environment», dans *Proceedings of the 18th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, p. 313–317. [10](#)
- Amour, L., S. Sami, S. Hoceini et A. Mellouk. 2015b, «An open source platform for perceived video quality evaluation», dans *Proceedings of the 11th ACM Symposium on QoS and Security for Wireless and Mobile Networks*, p. 139–140. [10](#)
- Amour, L., S. Sami, M. S. Mushtaq, S. Hoceini et A. Mellouk. 2016, «Perceived video quality evaluation based on interactive/repulsive relation between the qoe ifs», dans *2016 IEEE International Conference on Communications (ICC)*, IEEE, p. 1–7. [10](#)
- Amour, L., S. Souihi, S. Hoceini et A. Mellouk. 2015c, «A hierarchical classification model of qoe influence factors», dans *International Conference on Wired/Wireless Internet Communication*, Springer, Cham, p. 225–238. [10](#)
- Amour, L., S. Souihi, A. Mellouk et S. Mushtaq. 2020, «Q2abr : Qoe-aware adaptive video bit rate solution», *International Journal of Communication Systems*, vol. 33, n° 10, p. e4204. [10](#)
- Amour, L., V. Tong, S. Souihi, H. A. Tran et A. Mellouk. 2019, «Quality estimation framework for encrypted traffic (q2et)», dans *2019 IEEE Global Communications Conference (GLOBECOM)*, IEEE, p. 1–6. [10](#), [11](#)
- Balachandran, T., T. Abreu, M. Naloufi, S. Souihi, F. Lucas et A. Janne. 2022, «IoT and transfer learning based urban river quality prediction», dans *2022 IEEE Global Communications Conference : IoT and Sensor Networks (Globecom 2022 IoTSN)*, Rio de Janeiro, Brazil. [12](#)
- Bannour, F., S. Souihi et A. Mellouk. 2017a, «Distributed sdn control : Survey, taxonomy, and challenges», *IEEE Communications Surveys & Tutorials*, vol. 20, n° 1, p. 333–354. [11](#)
- Bannour, F., S. Souihi et A. Mellouk. 2017b, «Scalability and reliability aware sdn controller placement strategies», dans *2017 13th International Conference on Network and Service Management (CNSM)*, IEEE, p. 1–4. [11](#)
- Bannour, F., S. Souihi et A. Mellouk. 2018, «Adaptive state consistency for distributed onos controllers», dans *2018 IEEE Global Communications Conference (Globecom)*, IEEE, p. 1–7. [11](#)
- Bannour, F., S. Souihi et A. Mellouk. 2019, «Adaptive quorum-inspired sla-aware consistency for distributed sdn controllers», dans *2019 15th International Conference on Network and Service Management (CNSM)*, IEEE, p. 1–7. [11](#)
- Bannour, F., S. Souihi et A. Mellouk. 2020, «Adaptive distributed sdn controllers : application to content-centric delivery networks», *Future Generation Computer Systems*, vol. 113, p. 78–93. [11](#)
- Bennaceur, J., S. Souihi, H. Idoudi, L. A. Saldane et A. Mellouk. 2017, «Game-based secure sensing for the mobile cognitive radio network», dans *2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, IEEE, p. 1–7. [11](#)

- Chu, T.-T., V. Tong, H. A. Tran, S. Souihi, D. Q. Tran et A. Mellouk. 2019, «Nextlab : A new hybrid testbed and development platform for software-defined networking», dans *Proceedings of the Tenth International Symposium on Information and Communication Technology*, p. 186–190. [11](#), [12](#)
- Do, Q. H., S. Souihi, V. Tong, H. A. Tran et S. Mhadhbi. 2022, «How tezos blockchain can meet IoT?», dans *2022 IEEE Global Communications Conference : Next-Generation Networking and Internet (Globecom 2022 NGNI)*, Rio de Janeiro, Brazil. [12](#)
- Hoang, N.-T., H.-N. Nguyen, H.-A. Tran et S. Souihi. 2022, «A novel adaptive east&west interface for a heterogeneous and distributed sdn network», *Electronics*, vol. 11, n° 7, doi :10.3390/electronics11070975, ISSN 2079-9292. URL <https://www.mdpi.com/2079-9292/11/7/975>. [12](#)
- Nam Nguyen, H., H. Anh Tran, S. Fowler et S. Souihi. 2021, «A survey of blockchain technologies applied to software-defined networking : Research challenges and solutions», *IET Wireless Sensor Systems*, vol. 11, n° 6, doi :<https://doi.org/10.1049/wss2.12031>, p. 233–247. URL <https://ietresearch.onlinelibrary.wiley.com/doi/abs/10.1049/wss2.12031>. [12](#)
- Nguyen, H. N., S. Souihi, H. A. Tran et S. Fowler. 2022a, «A blockchain-based SDN East/West interface», dans *2022 IEEE Global Communications Conference : Next-Generation Networking and Internet (Globecom 2022 NGNI)*, Rio de Janeiro, Brazil. [12](#)
- Nguyen, T. A. M., S. Bannour Souihi, H. A. Tran et S. Souihi. 2022b, «When NLP meets SDN : an application to global internet exchange network», dans *2022 IEEE International Conference on Communications (ICC) : Next-Generation Networking and Internet Symposium (IEEE ICC'22 - NGNI Symposium)*, Seoul, Korea (South). [12](#)
- Tong, V., S. Souihi, H. A. Tran et A. Mellouk. 2021a, «Machine learning based root cause analysis for sdn network», dans *2021 IEEE Global Communications Conference (GLOBECOM)*, IEEE, p. 1–6. [11](#)
- Tong, V., S. Souihi, H. A. Tran et A. Mellouk. 2021b, «Sdn-based application-aware segment routing for large-scale network», *IEEE Systems Journal*. [11](#)
- Tong, V., H. A. Tran, S. Souihi et A. Mellouk. 2018, «A novel quic traffic classifier based on convolutional neural networks», dans *2018 IEEE Global Communications Conference (GLOBECOM)*, IEEE, p. 1–6. [11](#)
- Van, T., S. Souihi, H. A. Tran et A. Mellouk. 2020, «Service-centric segment routing mechanism using reinforcement learning for encrypted traffic», dans *2020 16th International Conference on Network and Service Management (CNSM)*, IEEE, p. 1–5. [11](#)
- Van, T., H. A. Tran, S. Souihi et A. Mellouk. 2018a, «Empirical study for dynamic adaptive video streaming service based on google transport quic protocol», dans *2018 IEEE 43rd Conference on Local Computer Networks (LCN)*, IEEE, p. 343–350. [11](#)
- Van, T., H. A. Tran, S. Souihi et A. MELLOUK. 2018b, «Network troubleshooting : survey, taxonomy and challenges», dans *2018 International Conference on Smart Communications in Network Technologies (SaCoNeT)*, IEEE, p. 165–170. [11](#)

Chapitre 3

Le paradigme SDN, le chaînant manquant

« *“The network is in my way”* »

James Hamilton, Amazon

Resumé

Dans ce chapitre, on s'intéressera au socle technologique pour la mise en oeuvre d'un réseau intelligent. Dans ce contexte, on pose le paradigme « Réseau défini par le logiciel », communément appelé le SDN (Software Defined Networking), comme l'une des briques de base pour atteindre cet objectif et on se questionne sur les apports et limites d'une telle architecture. Plus précisément, **nous proposons de nouvelles approches pour décentraliser le plan de contrôle SDN dans les réseaux à large échelle** tout en abordant certains des principaux problèmes associés.

Sommaire

3.1 Introduction et contexte	24
3.2 Motivation et état de l'art	24
3.3 contribution 1 : placement des contrôleurs	27
3.3.1 Formulation du problème	27
3.3.2 Approches proposées	29
3.3.3 Evaluation des performances	31
3.4 contribution 2 : modèle de cohérence continue	35
3.4.1 Mise en oeuvre	36
3.5 Validation expérimentale	37
3.5.1 Configuration expérimentale	37
3.6 contribution 3 : modèle de cohérence adaptatif	39
3.6.1 Approche proposée	40
3.6.2 Validation expérimentale	42
3.7 Conclusion et travail réalisé	45
3.8 Synthèses	47
3.8.1 Difficultés rencontrées	47
3.8.2 Forces et faiblesses	47
3.8.3 Possible extension	47
3.8.4 Impact potentiel des travaux	47
3.8.5 Retour d'expérience	47

3.1 Introduction et contexte

Afin de répondre aux besoins actuels et futurs, le réseau doit évoluer vers plus d'intelligence. Cette intelligence, basée sur un plan de connaissance global est censée apporter certains avantages et faciliter l'automatisation. Cependant, même si nous avons assisté à l'émergence de travaux allant dans ce sens, nous sommes encore loin du paradigme introduit par [Clark et collab. \[2003\]](#). Ces efforts se divisent en deux écoles, i) ceux qui proposent une transition douce et ii) ceux qui préconisent une approche plus radicale. Ces derniers soutiennent qu'une nouvelle architecture de réseau doit être construite à partir de zéro en rompant avec l'architecture conventionnelle et en apportant des modifications fondamentales pour répondre aux exigences actuelles et futures. Les partisans d'une transition douce, proposent des architectures qui continueraient à fonctionner dans le contexte actuel et donc à être compatibles avec les réseaux conventionnels tout en introduisant petit à petit les changements architecturaux nécessaires. Le paradigme récent du réseau piloté par logiciel (SDN) [[Feamster et collab., 2014](#)] offre la possibilité de programmer le réseau et facilite ainsi l'introduction d'approches de contrôle automatique et adaptatif en séparant le matériel (plan de données) et le logiciel (plan de contrôle), permettant leur évolution indépendante. Bien que répondant aux limites des architectures traditionnelles des réseaux, les architectures SDN centralisées soulèvent de nombreux défis d'évolutivité et de fiabilité. En effet, malgré l'intérêt que suscite le SDN, son déploiement dans le contexte industriel en est encore à ses débuts. Il faudra peut-être encore beaucoup de temps avant que la technologie ne mûrisse et que les efforts de normalisation portent leurs fruits pour que le potentiel du SDN soit pleinement exploité. Pour y répondre, il est nécessaire de faire évoluer l'architecture SDN vers une approche de systèmes physiquement distribués, mais logiquement centralisés. Néanmoins, il faut lever les verrous inhérents à certains cas d'application.

3.2 Motivation et état de l'art

L'initiative SDN, menée par l'Open Networking Foundation (ONF) [[ONF](#)], propose une nouvelle architecture ouverte pour relever les défis actuels des réseaux. Contrairement à l'architecture classique (figure 3.1) où les dispositifs de réseau sont intégrés verticalement et regroupent le logiciel et le matériel, l'architecture SDN (figure 3.1) élève le niveau d'abstraction en séparant les plans de données du contrôle du réseau. De cette façon, les dispositifs de réseau deviennent de simples commutateurs d'acheminement, tandis que toute la logique de contrôle est centralisée dans des contrôleurs logiciels fournissant un cadre de programmation flexible pour le développement d'applications spécialisées et pour le déploiement de nouveaux services.

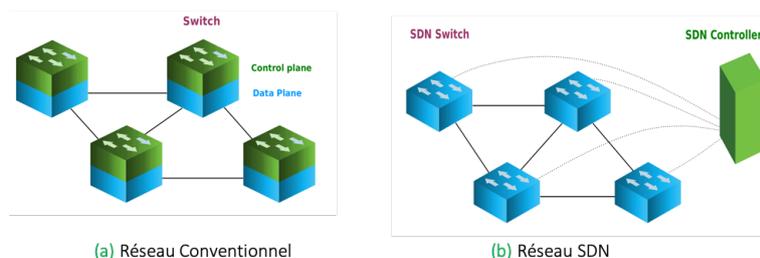


FIGURE 3.1 – Réseaux conventionnels vs SDN

L'architecture SDN est divisée verticalement en trois couches séparées par des inter-

faces (voir la figure 3.2). L'interface Nord ou « northbound », permet au plan de contrôle de communiquer avec le niveau supérieur. Inversement, l'interface sud ou « southbound », lui permet de communiquer avec le niveau inférieur. À ces deux interfaces s'ajoutent également les interfaces Est et Ouest, dont l'objectif est de permettre à différents contrôleurs de communiquer entre-eux.

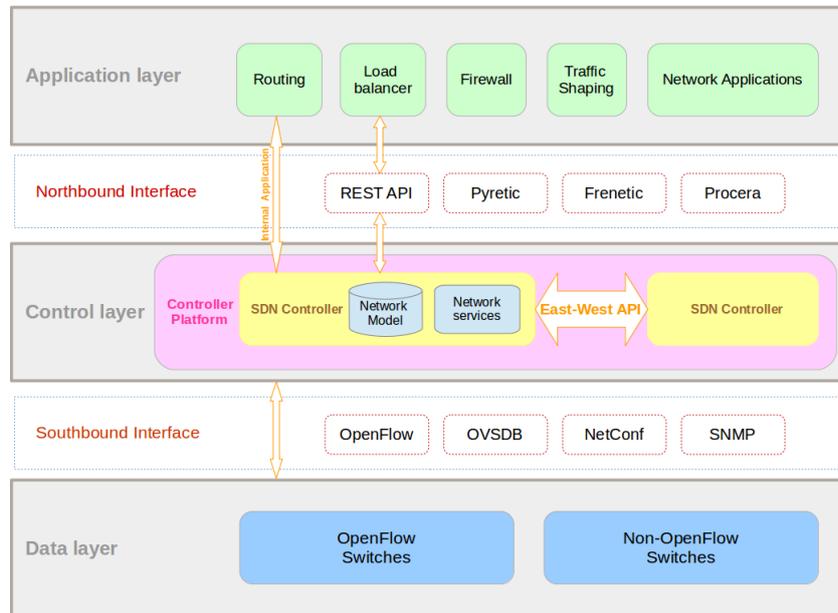


FIGURE 3.2 – Architecture des réseaux SDN

- Le plan de données, également appelé plan d'acheminement, consiste en un ensemble distribué d'éléments de réseau (principalement des commutateurs) chargés d'acheminer les paquets.
- le plan de contrôle consiste en un contrôleur logiciel centralisé chargé de gérer les communications entre les applications et les périphériques du réseau.
- Le plan "application" comprend les applications destinées à mettre en œuvre la logique et les stratégies de contrôle du réseau.

Bien que les SDN aient suscité beaucoup d'intérêt de la part de la communauté, il est à noter que les problématiques autour du contrôle distribué dans les réseaux SDN sont beaucoup moins étudiées (voir fig. 3.3).

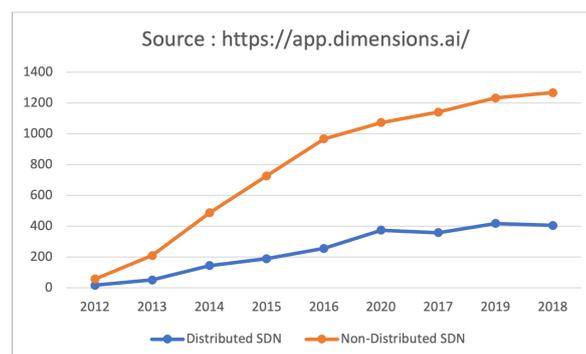


FIGURE 3.3 – Évolution de l'état de l'art sur les SDNs (en nombre de publications)

Dans ce contexte, nous avons proposé une taxonomie des plates-formes de contrôleurs SDN, les plus actuelles, en les classant de deux manières différentes : une classification physique et une classification logique [Bannour et collab., 2018]. Nous avons mené une analyse des principaux défis encore ouverts auxquels sont confrontés les plates-formes de contrôleurs SDN distribués existantes. Ces défis incluent les problèmes d'évolutivité, de fiabilité, de cohérence et d'interopérabilité du plan de contrôle SDN. En outre, nous avons exploré les approches potentielles pour relever ces défis en vue d'un déploiement optimal du SDN, et nous avons fourni des informations utiles sur les tendances émergentes et futures dans la conception de plans de contrôle SDN distribués plus efficaces. Le tableau 3.1 présente un bref résumé des principales caractéristiques et des indicateurs clés de performance des contrôleurs SDN examinés.

	Architecture du plan de contrôle	Conception du plan de contrôle	Langage de programmation	Scalabilité	Fiabilité	Cohérence
NOX [20]	Physiquement centralisée	–	C++	Très Limitée	Limitée	Forte
POX [1]	Physiquement centralisée	–	Python	Très Limitée	Limitée	Forte
Floodlight [2]	Physiquement centralisée	–	Java	Très Limitée	Limitée	Forte
SMArtLight [12]	Physiquement centralisée	–	Java	Très Limitée	Très Bonne	Forte
Ravana [27]	Physiquement centralisée	–	Python	Limitée	Très Bonne	Forte
ONIX [30]	Physiquement distribuée Logiquement centralisée	Plat	Python C	Très Bonne	Bonne	Faible Forte
HyperFlow [41]	Physiquement distribuée Logiquement centralisée	Plat	C++	Bonne	Bonne	Éventuelle
ONOS [11]	Physiquement distribuée Logiquement centralisée	Plat	Java	Très Bonne	Bonne	Faible Forte
OpenDayLight [3]	Physiquement distribuée Logiquement centralisée	Plat	Java	Très Bonne	Bonne	Forte
B4 [26]	Physiquement distribuée Logiquement centralisée	Hierarchique	Python C	Bonne	Bonne	N/A
Kandoo [22]	Physiquement distribuée Logiquement centralisée	Hierarchique	C C++ Python	Très Bonne	Limitée	N/A
DISCO [37]	Physiquement distribuée Logiquement distribuée	Plat	Java	Bonne	Limitée	Forte (inter-domaines)
SDX [21]	Physiquement distribuée Logiquement distribuée	Plat	Python	Limitée	N/A	Forte
DevoFlow [15]	Physiquement distribuée Logiquement centralisée	N/A	Java	Bonne	N/A	N/A
DIFANE [44]	Physiquement distribuée Logiquement centralisée	N/A	–	Bonne	N/A	N/A

TABLEAU 3.1 – Principales caractéristiques des contrôleurs SDN existants

Les contrôleurs physiquement centralisés tels que NOX [Gude et collab., 2008], POX [pox] et FloodLight [flo, 2015] souffrent de problèmes d'évolutivité et de fiabilité. Des solutions comme DevoFlow [Curtis et collab., 2011] et DIFANE [Yu et collab., 2010] ont tenté de résoudre ces problèmes d'évolutivité en repensant le protocole OpenFlow. Certaines propositions de SDN distribués telles que Kandoo [Hassas Yeganeh et Ganjali, 2012], préconisent une organisation hiérarchique du plan de contrôle pour améliorer encore

l'évolutivité là où d'autres alternatives optent pour une organisation plate afin d'améliorer la fiabilité et les performances (particulièrement la latence). D'autre part, les plateformes distribuées comme Onix [Koponen et collab., 2010], HyperFlow [Koponen et collab., 2010], ONOS [Berde et collab., 2014] et OpenDaylight [ope, 2016], se sont concentrées sur la construction de nouveaux modèles de cohérence. Onix [Koponen et collab., 2010] a par exemple choisi les DHT et les bases de données transactionnelles pour la distribution de l'état du réseau plutôt que le système Publish/Subscribe utilisé par HyperFlow [Toonchian et Ganjali, 2010]. Une autre classe de solutions différentes a été récemment introduite par DISCO qui a promu un plan de contrôle logiquement distribué basé sur les AS existants dans l'Internet.

À la lumière de l'étude que nous avons menée, il est évident qu'il existe différentes approches pour construire une architecture SDN distribuée. Certaines de ces approches répondent mieux que d'autres à certains critères de performance, mais échouent sur d'autres aspects et il est évident qu'aucune des plateformes de contrôle SDN proposées n'a répondu aux challenges discutés et rempli tous les indicateurs clés de performance requis pour un déploiement optimal du SDN. Dans notre cas et dans le respect d'une vision pragmatique de la recherche, nous proposons d'aborder le problème du contrôle SDN distribué en nous concentrant sur deux défis majeurs gérables qui, bien que corrélés, pourraient être traités comme des problèmes de recherche distincts : le problème du placement des contrôleurs (1) et le problème de la diffusion des connaissances (cohérence) (2). L'objectif est de proposer un plan de contrôle SDN distribué, évolutif et fiable. Le premier problème concerne le nombre requis de contrôleurs SDN ainsi que leur emplacement approprié par rapport aux objectifs souhaités. Le second problème concerne le type et la quantité d'informations réseau à partager entre les instances de contrôleurs SDN, compte tenu du niveau souhaité de cohérence et de performance de l'état des applications. Dans la suite de ce chapitre nous allons détailler ces contributions.

3.3 contribution 1 : placement des contrôleurs

Concevoir un plan de contrôle SDN distribué (mais logiquement centralisé) évolutif et fiable dépend essentiellement du placement de ces contrôleurs SDN. Plus précisément, le *problème de placement des contrôleurs* [Das et collab., 2019] consiste à trouver le *nombre requis* et *l'emplacement* approprié des contrôleurs SDN (parmi les nœuds du réseau) afin d'obtenir le meilleur compromis entre les mesures de performance et de fiabilité.

3.3.1 Formulation du problème

On considère le réseau comme un graphe $G = (V, E)$; où V représente les nœuds du réseau comprenant les contrôleurs et les commutateurs tandis que l'ensemble des arêtes E représente les liens reliant ces nœuds du réseau. Le poids des arêtes représente les performances des liens (par exemple les latences entre chaque paire de nœuds). $d(s, c)$ et la latence entre un nœud de commutation $s \in V$ et un nœud de contrôleur $c \in V$. Ces informations sont stockées dans la base de connaissance en charge de maintenir une vue de la topologie

Nous formulons le problème de placement du contrôleur comme un problème d'optimisation multi-objectif en fonction d'un ensemble de mesures de performance et de fiabilité.

Métriques de performance

L'optimisation des performances du plan de contrôle est d'une importance capitale dans les réseaux à grande échelle avec des exigences strictes en matière de temps de réponse et où des délais de propagation élevés peuvent conduire à des comportements incohérents et incorrects des services réseau.

En particulier, la latence moyenne et la latence maximale entre les commutateurs et leurs contrôleurs associés sont deux métriques introduites pour la première fois par [Heller et collab. \[2012\]](#) afin de quantifier la performance du placement. La latence moyenne évalue la qualité globale du placement du contrôleur sans égard aux cas particuliers (latence élevée entre les commutateurs et leurs contrôleurs associés). La latence maximale est une métrique à diminuer car elle représente les cas particulier à éviter, par exemple, un commutateur isolé. Nous avons défini ces 2 métrique comme suit :

- Latence moyenne entre le commutateur et le contrôleur :

$$\pi^{Avg-s2c-Latence}(C) = \frac{1}{n} \sum_{(s \in S)} \min_{(c \in C)} d(s, c) \quad (3.1)$$

- Latence maximale entre le commutateur et le contrôleur :

$$\pi^{Max-s2c-Latence}(C) = \max_{(s \in S)} \min_{(c \in C)} d(s, c) \quad (3.2)$$

Par ailleurs, il est également intéressant de considérer les latences inter-contrôleurs dans le placement [[Lange et collab., 2015](#); [Killi et Rao, 2019](#)]. Les contrôleurs SDN physiquement distribués doivent en effet être placés aussi près que possible les uns des autres afin de réduire le coût du maintien de la cohérence, c'est-à-dire la communication entre contrôleurs en vue de leurs synchronisation. En conséquence, les emplacements des contrôleurs SDN doivent être sélectionnés de manière à minimiser les latences inter-contrôleurs moyennes et maximales définies comme suit :

- Latences inter-contrôleurs :

$$\pi^{Avg-c2c-Latence}(C) = \frac{1}{|C|} \sum_{(c_1, c_2 \in C)} d(c_1, c_2) \quad (3.3)$$

$$\pi^{Max-c2c-Latence}(C) = \max_{(c_1, c_2 \in C)} d(c_1, c_2) \quad (3.4)$$

La connaissance de la capacité du contrôleur est également un facteur de performance important qui doit être pris en compte dans le processus de placement afin d'éviter le risque de surcharge du contrôleur et, par conséquent, les problèmes de performance qui y sont liés (par exemple, les retards induits au niveau du contrôleur). Une façon d'aborder l'aspect de la surcharge des contrôleurs est de supposer que tous les contrôleurs ont des capacités égales (en termes de nombre de nœuds contrôlés) et de garantir une distribution égale de la charge du plan de contrôle. Par exemple, chaque contrôleur peut être chargé à 80% de sa pleine capacité, laissant une marge de capacité de contrôleur de 20% pour éviter une surcharge occasionnelle du contrôleur. Un autre schéma d'équilibrage de la charge des contrôleurs consiste à introduire une métrique supplémentaire de déséquilibre de la charge à minimiser. Cette métrique représente la différence entre le nombre maximum et minimum de nœuds de réseau assignés à un contrôleur pour un placement donné et elle est définie comme suit :

$$\pi^{Déséquilibre\ de\ charge}(C) = \max_{(c \in C)} n_c - \min_{(c \in C)} n_c \quad (3.5)$$

Métriques de fiabilité

La fiabilité du réseau est également cruciale pour les réseaux SDN opérationnels. Il convient donc d'accorder une attention particulière aux métriques lors de la conception de plateformes SDN distribuées. La fiabilité des SDN peut être assurée en plaçant les contrôleurs de manière à atténuer l'impact de leurs défaillances. Un mécanisme commun de récupération après une panne du contrôleur principal consiste à affecter les commutateurs orphelins aux contrôleurs les plus proches. Néanmoins, les exigences en matière de temps de réponse doivent être respectées. Nous pouvons utiliser la métrique de latence maximale (à minimiser). En effet, le pire scénario pour un commutateur réseau serait la défaillance simultanée des $(k - 1)$ contrôleurs SDN les plus proches. L'atténuation de ce scénario de défaillance du plan de contrôle implique de minimiser les latences entre les commutateurs s et leurs contrôleurs fonctionnels respectifs les plus éloignés $C_{Fu}(s)$ comme suit :

$$\pi_{F(k-1)}^{Max-s2c-Latence}(C) = \max_{(s \in S, c \in C_{Fu}(s))} d(s, c) \quad (3.6)$$

En pratique, il est plus courant que les défaillances des contrôleurs primaires se produisent un par un. Ainsi, la réduction de l'impact de ce scénario de défaillance du contrôleur implique de minimiser le maximum de latences entre les commutateurs s et leurs contrôleurs respectifs les plus proches $C_{Cl}(s)$, comme exprimé ci-dessous :

$$\pi_{F(1)}^{Max-s2c-Latence}(C) = \max_{(s \in S, c \in C_{Cl}(s))} d(s, c) \quad (3.7)$$

Il est donc important de mettre en place des stratégies prenant en compte plusieurs critères d'évolutivité et de fiabilité pour le *placement de contrôleurs* SDN distribués à large échelle à l'aide de différents types d'algorithmes d'optimisation multi-critères.

3.3.2 Approches proposées

Toujours dans une vision pragmatique de la recherche, nous nous sommes posés la question de savoir si le problème de placement de contrôleurs SDN pouvait être considéré comme un problème de classification. En effet, ce type de problème est bien connu et peut être résolu plus au moins efficacement par des heuristiques peu complexes. Nous avons donc proposé une approche de placement de contrôleurs sensibles à de multiples critères de fiabilité et de performance. Cette solution, nommée PAM-B est une évolution de l'approche PAM (Partitioning Around Medoids) [Ganatra, 2012]. À des fins de comparaison, nous l'avons comparé à un algorithme génétique modifié appelé NSGA-II (Non-dominated Sorting Genetic Algorithm II) [Ahmadi et collab., 2015].

PAM (Partitioning Around Medoids) est une méthode de k -Medoids qui partitionne l'ensemble de données de N objets (nœuds) en k clusters représentés par k Medoids (contrôleurs). L'idée de PAM est de trouver l'ensemble des Medoids qui améliorent la qualité globale du clustering qui est mesurée sur la base de la dissimilarité moyenne de tous les objets par rapport à leur Medoid le plus proche. Dans notre cas, nos métriques sont d'importance égale, ce qui fait que la fonction de dissimilarité D pour un placement donné $C \in CP$ (les configurations de placement) est calculée comme la somme normalisée de tous les objectifs pondérés O avec les poids associés égaux à $\frac{1}{M}$:

$$D^{PAM-B}(C) = \sum_{i \in M} \left(\frac{1}{M} \right) \times N(O_i) \quad (3.8)$$

$$avec : N(O_i) = \frac{O_i(C) - \min_{(C \in CP)} O_i(C)}{\max_{(C \in CP)} O_i(C) - \min_{(C \in CP)} O_i(C)}$$

NSGA-II est un algorithme génétique rapide et populaire pour l'optimisation multi-objectif. En plus des opérateurs génétiques classiques (croisement et mutation), NSGA-II utilise d'autres mécanismes de classement (tri non dominé et distance d'encombrement) pour créer la population de solutions candidates de la prochaine génération. L'idée principale de NSGA-II est de faire évoluer cette population vers un ensemble de solutions non dominées (le front de Pareto) offrant les meilleurs compromis entre les objectifs considérés.

Ces approches multi-critères ont été comparées en termes de temps de calcul, et de qualité des configurations finales de placement des contrôleurs SDN dans un environnement applicatif et en considérant plusieurs stratégies [Bannour et collab., 2017].

Stratégie 1 : Mesures basées sur les performances

La stratégie 1 résout le problème de placement des contrôleurs SDN en se basant sur les métriques de performance liées à la latence indiquée dans (3.1) et (3.2) tout en suivant la méthode habituelle d'affectation contrôleur-commutateur du plus court chemin. Nous adoptons également un schéma d'équilibrage de la charge utilisant la métrique de déséquilibre de la charge (3.5) proposée par Hock et collab. [2013] et nous évaluons le risque de surcharge du contrôleur. En conséquence, la NSGA-II multi-objectif est lancée avec ces trois objectifs à minimiser tandis que le PAM-B minimise la fonction de dissimilarité représentée par la somme normalisée équilibrée des trois objectifs considérés conformément à (3.9) :

$$D_1^{\text{PAM-B}}(C) = \text{Avg} \left(\begin{aligned} &N(\pi^{\text{Avg-Latence}}(C)), \\ &N(\pi^{\text{Max-Latence}}(C)), \\ &N(\pi^{\text{Déséquilibre de charge}}(C)) \end{aligned} \right) \quad (3.9)$$

Stratégie 2 :

La stratégie 2 intègre, en plus des mesures de performance liées à la latence mentionnées précédemment, une contrainte sur la capacité des contrôleurs (80 % de leurs capacités). En même temps, cette stratégie assure une affectation équitable des commutateurs du réseau à leurs contrôleurs les plus proches faiblement chargés. Ainsi, dans la stratégie 2, PAM-B minimise la même fonction de dissimilarité utilisée dans la stratégie 1 en rajoutant des règles d'affectation.

Stratégie 3 :

En plus des mesures de performance de la stratégie 1, nous adoptons un autre schéma d'équilibrage de la charge en utilisant la mesure du déséquilibre de la charge et nous étudions le risque de surcharge du contrôleur.

La formule suivante définit la fonction de dissimilarité de PAM-B en fonction des trois objectifs considérés :

$$D_3^{\text{PAM-B}}(C) = \left(\frac{1}{3}\right) \times N(\pi^{\text{Latence Moyenne}}(C)) + \left(\frac{1}{3}\right) \times N(\pi^{\text{Latence maximale}}(C)) + \left(\frac{1}{3}\right) \times N(\pi^{\text{Déséquilibre de charge}}(C)) \quad (3.10)$$

Stratégie 4 :

La stratégie 4 fournit un modèle d'optimisation du placement des contrôleurs SDN qui inclut des mesures de fiabilité et des mesures de performance. En ce qui concerne les métriques de fiabilité, il est possible d'inclure une variante de métrique de fiabilité qui s'attaque au pire scénario de défaillance du contrôleur ou une variante qui s'attaque à un scénario de défaillance du contrôleur plus courant, en plus des métriques de performance précédentes.

Les fonctions de dissimilarité de PAM-B pour les deux variantes sont calculées selon les formules suivantes :

$$D_4^{\text{PAM-B}(k-1)}(C) = \left(\frac{1}{4}\right) \times N(\pi^{\text{LatenceMoyenne}}(C)) + \left(\frac{1}{4}\right) \times N(\pi^{\text{Latencemaximale}}(C)) \\ + \left(\frac{1}{4}\right) \times N(\pi^{\text{Déséquilibredecharge}}(C)) + \left(\frac{1}{4}\right) \times N(\pi_{F(k-1)}^{\text{Latencemaximale}}(C)) \quad (3.11)$$

$$D_4^{\text{PAM-B}(1)}(C) = \left(\frac{1}{4}\right) \times N(\pi^{\text{LatenceMoyenne}}(C)) + \left(\frac{1}{4}\right) \times N(\pi^{\text{Latencemaximale}}(C)) \\ + \left(\frac{1}{4}\right) \times N(\pi^{\text{Déséquilibredecharge}}(C)) + \left(\frac{1}{4}\right) \times N(\pi_{F(1)}^{\text{Latencemaximale}}(C)) \quad (3.12)$$

3.3.3 Evaluation des performances

Paramètres de simulation

Dans NSGA-II, le paramètre de simulation *nombre maximal d'évaluations de la fonction objectif* (MaxEvaluations) a été utilisé comme critère d'arrêt. Les valeurs de ce paramètre (de 20000 à 140000) dépendent du nombre d'objectifs impliqués dans chaque stratégie (3 à 4) et de la taille du réseau (de 20 à 1000 nœuds) dans chaque scénario.

Résultats de la simulation

Pour chaque stratégie, étant donné une topologie donnée, nous évaluons les solutions de placement de chaque algorithme. PAM-B génère une solution de regroupement basée sur la mesure de dissimilarité à pondération égale (3.8). De même, dans NSGA-II, nous considérons la solution de placement la plus juste par rapport à nos critères parmi les solutions Pareto optimales non dominées générées et représentant les compromis possibles entre les objectifs. Pour ce faire, nous sélectionnons la solution de Pareto S qui réduit le mieux l'écart total entre les valeurs des objectifs et leurs valeurs optimales respectives dans l'ensemble des solutions optimales de Pareto P. Cela correspond à la solution de Pareto ayant la valeur minimale de la mesure ci-dessous :

$$a(S) = \sum_{i \in M} \left(\frac{1}{M}\right) \times \frac{O_i(S) - \min_{(S \in P)} O_i(S)}{\max_{(S \in P)} O_i(S) - \min_{(S \in P)} O_i(S)} \quad (3.13)$$

Plusieurs scénarios de simulation sont réalisés en suivant les stratégies considérées et en utilisant divers types de topologies de taille différente; de **20 à 1000** nœuds. Cela nous a permis de comparer nos stratégies de placement, d'analyser les performances de nos algorithmes d'optimisation et d'étudier l'évolutivité de notre approche qui est principalement destinée aux déploiements à grande échelle.

Comme le montre la figure 3.4, PAM-B est meilleur que NSGA-II pour minimiser la latence moyenne (de l'ordre de 6,68 % en moyenne) et la latence maximale (de l'ordre de 10,53 % en moyenne) sur tous les scénarios. En plus de viser un certain niveau de performance en matière de latence, la stratégie 3 considère l'équilibrage de la charge comme un autre facteur de performance important. Cependant, la minimisation de la métrique de déséquilibre de charge n'élimine pas le risque de surcharge des contrôleurs d'où l'importance de considérer un seuil comme dans le cas de la stratégie 2.

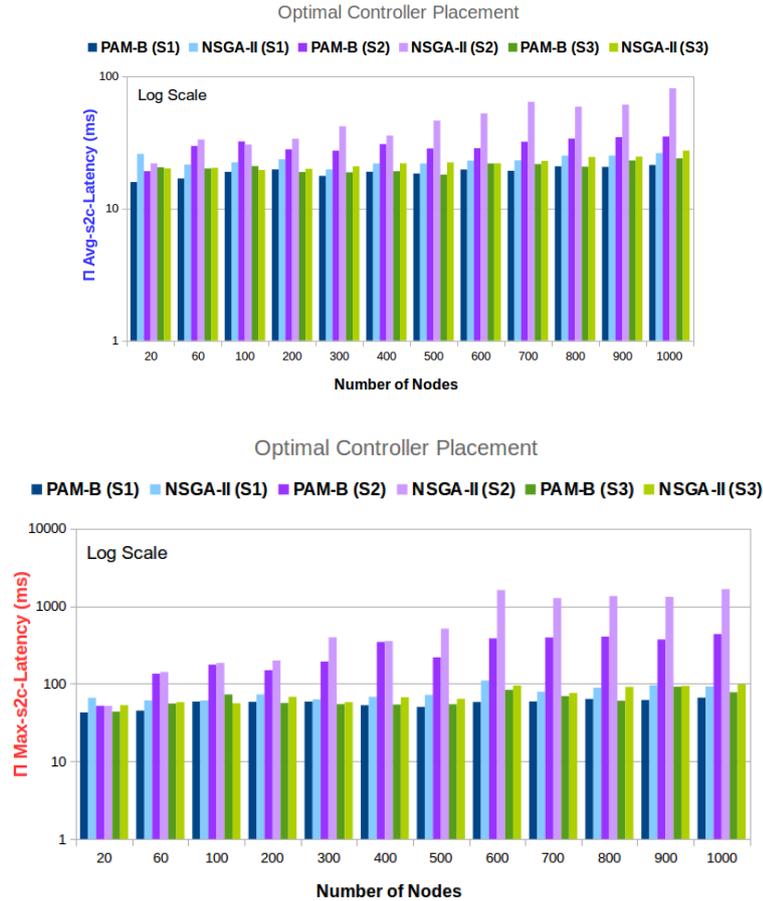


FIGURE 3.4 – Stratégie 1, 2 et 3 : Mesures de performance basées sur la latence

La stratégie 4 fait intervenir des mesures de fiabilité en plus des mesures de performance de la stratégie 1. La figure 3.5 compare, pour chaque variante considérée, les valeurs de la latence maximale dans le cas "sans défaillance" avec celle de la latence maximale dans le cas "avec défaillance". Elle montre que PAM-B et NSGA-II obtiennent des résultats assez similaires.

La figure 3.6 étudie le coût de performance de l'implication des critères de fiabilité dans l'optimisation du placement des contrôleurs. On constate que l'optimisation des mesures de fiabilité n'a pas eu d'impact important sur les mesures de performance telles que la latence maximale (dans le cas sans défaillance).

Discussion

L'évaluation de PAM-B et de NSGA-II sur nos stratégies a révélé que dans la plupart des scénarios, PAM-B surpasse NSGA-II en termes de qualité des solutions finales. PAM-B offre un compromis équilibré entre les métriques de performance et de fiabilité, ainsi que des résultats stables sur les deux stratégies, alors que les performances de NSGA-II par rapport à ces métriques sont parfois imprévisibles et dépendent de la stratégie. La comparaison des temps d'exécution entre PAM-B et NSGA-II a montré des tendances similaires sur les scénarios. Leurs complexités de calcul sont en effet proches et respectivement égales à $O(k(n - k)^2)$ (k est le nombre de médoides (contrôleurs) et n est le nombre d'objets (la taille du réseau)) et $O(MN^2)$ (M est le nombre d'objectifs et N est la taille de la population). Il convient de noter que le temps de calcul de PAM peut être

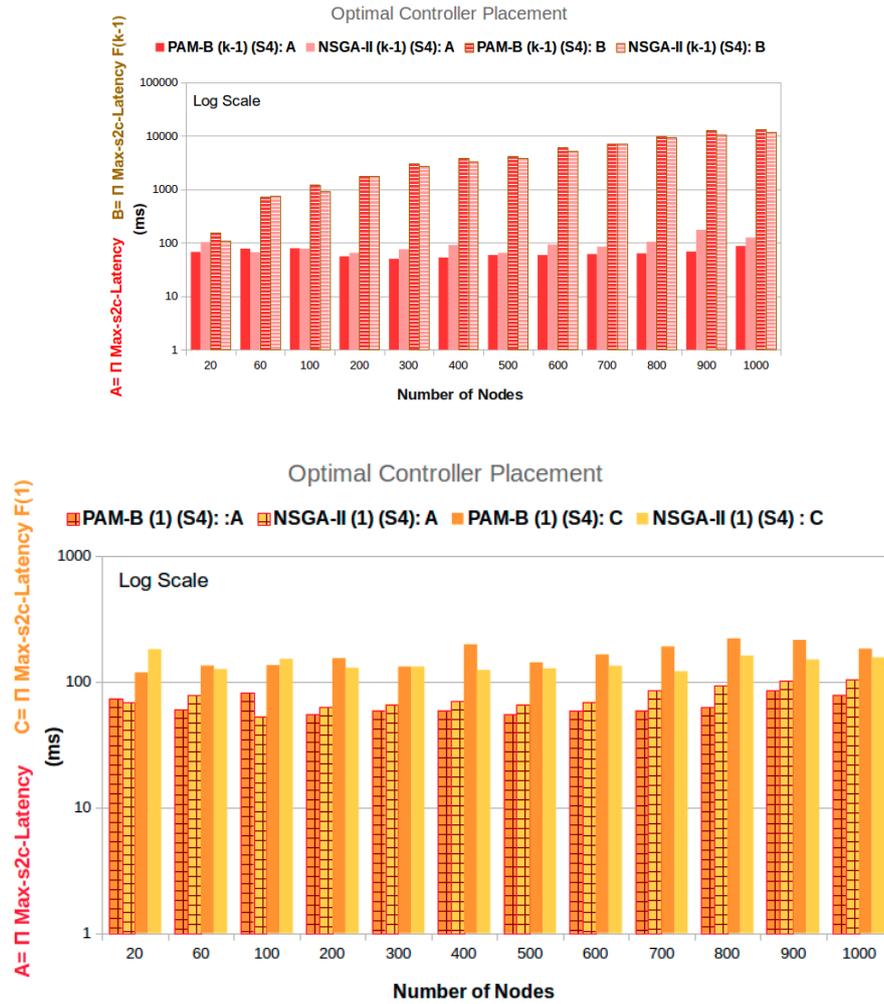


FIGURE 3.5 – Strategie 4 : Mesures de fiabilité : (latences maximales dans les scénarios sans défaillance)

amélioré. CLARA [Shraddha et Emmanuel, 2014], une variante de PAM basée sur l'échantillonnage, est recommandée pour les grands ensembles de données. Sa complexité est de $O(ks^2 + k(n - k))$, où k est le nombre de contrôleurs, n est la taille du réseau et s est la taille de l'échantillon.



FIGURE 3.6 – Strategie 4 : Mesures de la performance

3.4 contribution 2 : modèle de cohérence continue

Au delà de la problématique de placement des contrôleurs, l'adoption généralisée d'une architecture SDN physiquement distribuée fait face à des limites importantes en termes d'évolutivité, de disponibilité et de fiabilité, notamment dans le contexte de déploiements de réseaux réels à grande échelle. En fait, garantir un niveau approprié de cohérence de l'état des contrôleurs tout en maintenant de bonnes performances dans un réseau SDN est une tâche difficile, comme l'indique le théorème CAP. Il est donc nécessaire de considérer différents compromis entre les propriétés de cohérence et de performance.

Les implémentations actuelles des plateformes de contrôleurs SDN distribués [ono, ONOS; ope, 2016] utilisent des modèles de cohérence forte et/ou éventuelle pour leurs applications de contrôle internes et externes. Ces modèles présentent différents avantages et inconvénients.

Le modèle dit de cohérence forte assure une vue de l'état du réseau fortement cohérente pour garantir le comportement correct des applications. Dans le SDN, le modèle de cohérence forte garantit que tous les contrôleurs dans un même cluster ont accès aux informations les plus récentes à tout moment. Cependant, il introduit une surcharge de synchronisation entre contrôleurs et entrave le temps de réponse et le débit, ce qui a un impact sur la disponibilité et les performances du réseau. En fait, la cohérence forte repose sur un processus de synchronisation bloquant qui empêche les commutateurs de lire les données, à moins que les contrôleurs ne soient entièrement mis à jour, ce qui affecte la disponibilité et l'évolutivité du réseau. La cohérence forte est une exigence pour certaines applications qui privilégient les propriétés de cohérence et de correction à la disponibilité. Dans les plates-formes de contrôleurs actuelles, la cohérence forte est généralement obtenue en utilisant par exemple Paxos ou RAFT [Ongaro et Ousterhout, 2014].

D'un autre côté, le modèle de cohérence éventuelle permet une plus grande disponibilité, en particulier dans le contexte des réseaux SDN distribués à grande échelle prenant en charge des applications évolutives. Cependant, il peut présenter une vue du réseau temporairement incohérente qui conduit à des comportements applicatifs incorrects. De nombreuses applications optent pour la cohérence éventuelle afin de garantir la haute disponibilité et les performances à l'échelle. Les systèmes de bases de données modernes, notamment Dynamo [Sivasubramanian, 2012] et Cassandra [Lakshman et Malik, 2010], prennent en charge par défaut les paramètres de cohérence éventuelle favorisant une disponibilité (accès rapide aux données) et une évolutivité extrêmement élevées.

Comme première réponse à ce dilemme, nous avons introduit l'utilisation d'un modèle de cohérence multi-niveaux adaptatif suivant le concept de cohérence continue [Yu et Vahdat, 2000]. Ce modèle présente de nombreux avantages par rapport aux extrêmes que sont la cohérence forte et la cohérence éventuelle, en particulier dans les déploiements à grande échelle. Contrairement aux approches de cohérence statique, les modèles à cohérence adaptative ajustent leur niveau de cohérence au moment de l'exécution pour atteindre les exigences souhaitées en matière de performance et de cohérence des applications. Cette alternative offre de nombreux avantages : Elle épargne aux concepteurs d'applications la tâche de développer des applications complexes nécessitant la mise en œuvre de plusieurs modèles de cohérence, elle offre aux applications une certaine robustesse face aux conditions soudaines du réseau et elle réduit la surcharge de la distribution d'état entre les contrôleurs sans compromettre les performances de l'application [Aslan et Matrawy, 2016]. Notre stratégie d'adaptation de la cohérence consiste principalement à ajuster continuellement le niveau de cohérence en fonction des incohérences

d'état observées par rapport aux exigences de l'application (SLA) dans des conditions de réseau changeantes. Il se présente comme un modèle de cohérence adaptative utilisant *la période de réconciliation Anti-Entropie* comme outil de contrôle pour un réglage fin et adaptatif des niveaux de cohérence. Cette approche a été implémentée pour une application de routage à la source au-dessus des contrôleurs ONOS. L'objectif est également de minimiser la surcharge réseau de réconciliation Anti-Entropie par rapport à la stratégie de cohérence statique d'ONOS. L'approche mise en oeuvre se base sur un middleware appelé TACT [Yu et Vahdat, 2000] qui propose de régler de manière adaptative les exigences de cohérence et de disponibilité des services en suivant le concept de cohérence continue. Le modèle de cohérence continue explore ainsi l'espace sémantique entre ces deux types de modèles traditionnels : Il offre un continuum de modèles de cohérence intermédiaires. Ces degrés de cohérence quantifiables peuvent être exploités par les applications pour explorer, au moment de l'exécution, leurs propres compromis entre cohérence et disponibilité, tout en tenant compte de l'évolution du réseau et des conditions de service. Plus précisément, TACT limite la quantité d'incohérence et de divergence entre les répliques d'une manière spécifique à l'application. Fondamentalement, les applications spécifient leur sémantique de cohérence à travers les *conits*; un ensemble de mesures qui capturent le spectre de cohérence : *L'erreur numérique, L'erreur d'ordre, Obsolescence*.

Ainsi, pour chaque conit, la cohérence est quantifiée de manière continue le long d'un vecteur tridimensionnel :

$$\text{Cohérence} = (\text{Erreur numérique}, \text{Erreur d'ordre}, \text{Obsolescence}) \quad (3.14)$$

- *L'erreur numérique* limite l'écart entre la valeur fournie au client de l'application et la valeur "finale" la plus cohérente.
- *L'erreur d'ordre* limite l'incohérence par le nombre d'écritures provisoires pour chaque réplique.
- *L'obsolescence* place une limite en temps réel sur le délai de propagation des écritures entre les répliques.

3.4.1 Mise en oeuvre

D'un point de vue pratique, nous avons fait évoluer le modèle de cohérence éventuelle du contrôleur ONOS en conservant la technique de réplification optimiste utilisée pour répliquer les événements et les mises à jour entre les contrôleurs et en agissant sur le processus d'anti-entropie. En effet ce processus est naturellement exécuté sur chaque contrôleur périodiquement à des intervalles fixes (toutes les 3-5 secondes) (*voluntary anti-entropy*) causant trop de surcharge et affectant l'évolutivité et la performance du système. Nous avons donc fait évoluer ce mécanisme afin que le processus d'anti-entropie soit programmé uniquement lorsque la cohérence du système est en danger (*compulsory anti-entropy*). En d'autres termes, le choix de la période de réconciliation anti-entropie pour chaque réplique de contrôleur (*cohérence par réplique*) est basé sur la correction du système par rapport aux exigences de cohérence exprimées par les applications. Ainsi, à chaque contrôleur et pour chaque état de l'application, le niveau de cohérence est adapté dynamiquement en fonction des valeurs calculées des métriques de cohérence. Nous avons considéré une application SDN de routage à la source comme cas d'application. Cette application est basée sur le calcul des coûts du plus court chemin entre les hôtes source et destination. Puisque l'état du graphe de topologie est traité de manière éventuellement cohérente, l'état de l'application est également considéré comme éventuellement cohérent.

Dans notre application de routage f , le chemin entre l'hôte source A et l'hôte de destination B peut être défini comme un *conit*. En outre, nous soutenons qu'une exigence de cohérence importante pour notre application de contrôle est l'optimalité du résultat du coût de calcul du chemin instantané (en termes de nombre de sauts dans notre cas) qui est capturé par la métrique *d'Erreur*

numérique. L'erreur numérique de notre conit C peut être définie comme la différence relative entre la valeur du coût du "plus court chemin" x_{local} telle que perçue par une réplique locale, et sa valeur "optimale" $x_{optimal}$ au niveau d'une réplique qui a atteint un certain état cohérent "absolue". Cette erreur est continuellement limitée au moment de l'exécution à l'aide d'un seuil $T(f)$ défini par l'application (en pourcentage) comme suit :

$$Erreurnumérique(C_f) = \left(\frac{|x_{local} - x_{optimal}|}{x_{optimal}} \right) < T(f) \quad (3.15)$$

Enfin, il convient de noter que d'autres sémantiques de cohérence pour l'application de routage à la source pourraient être exprimées à l'aide de *l'obsolescence* et de *l'erreur d'ordre*.

En fait, en cas de défaillance d'un contrôleur, les autres contrôleurs du cluster qui détectent cette défaillance conservent une capture de leur propre graphe topologique au moment de la défaillance. Pendant le contrôle de cohérence périodique, ils utilisent ce graphe topologique stocké pour estimer l'incohérence du contrôleur défaillant, qui est égale à la différence relative entre le coût du plus court chemin "local" (calculé sur la base de l'état actuel du graphe topologique) et le coût du plus court chemin tel qu'il est perçu par le contrôleur après le rétablissement. Lorsque le contrôleur défaillant se rétablit, les autres contrôleurs prennent une décision anti-entropie en fonction de l'erreur numérique vérifiée. Si l'erreur dépasse un seuil de cohérence "alarmant" fixé à l'avance par l'application, un processus d'anti-entropie est lancé pour réparer l'état du contrôleur défaillant. Dans le cas contraire, l'incohérence est considérée comme tolérée par l'application et une session d'anti-entropie peut être programmée par la suite si l'état du contrôleur dérive de manière significative.

3.5 Validation expérimentale

3.5.1 Configuration expérimentale

Nos expériences sont réalisées sur un serveur Ubuntu 16.04 LTS utilisant ONOS 1.13. Nous utilisons également Mininet 2.2.1. Wireshark est utilisé comme sonde pour capturer le trafic entre contrôleurs (port TCP 9876).

Pour valider notre approche, que nous appellerons ONOS-WAC (ONOS-With Adaptive Consistency), nous avons envisagé de nombreux scénarios de test. Dans chaque scénario, nous exécutons un cluster de N instances de contrôleur ONOS, contrôlant une topologie de réseau Mininet de S commutateurs.

Afin de créer des incohérences d'état entre les instances de contrôleur dans le cluster, nous créons différents scénarios de panne de contrôleur F. Peu après une panne d'un contrôleur dans un scénario spécifique S_i , nous envisageons de modifier la topologie du réseau en mettant hors service les commutateurs et les liens du réseau le long du chemin le plus court (calculé par l'application) entre l'hôte source A et l'hôte de destination B. Ainsi, après la récupération, le contrôleur aura une vue incohérente de la topologie du réseau par rapport au reste des contrôleurs du cluster. Selon l'approche que nous proposons, cette incohérence dans la vue de la topologie du réseau affecte l'optimalité du calcul du coût du plus court chemin utilisé par l'application de routage à la source qui s'exécute sur le contrôleur récupéré. *L'Erreur numérique* induite est susceptible de déclencher un processus de synchronisation réalisé par des tâches anti-entropie.

La figure 3.7 illustre le trafic entre contrôleurs capturé pendant la période du scénario de test dans un cluster ONOS et dans un cluster ONOS-WAC. Par exemple, le premier pic de trafic dans les deux figures (à $t = 90s$) correspond à une commande CLI Mininet "pingall" exécutée pour la découverte de la topologie. À $t = 150s$, nous simulons un scénario de panne en mettant hors service une instance de contrôleur. Cette action est suivie d'autres changements de topologie (par exemple, des liens hors service) correspondant aux pics suivants dans les deux figures. À $t = 180s$, nous remettons en service le contrôleur défaillant, ce qui entraîne un pic de trafic qui semble plus

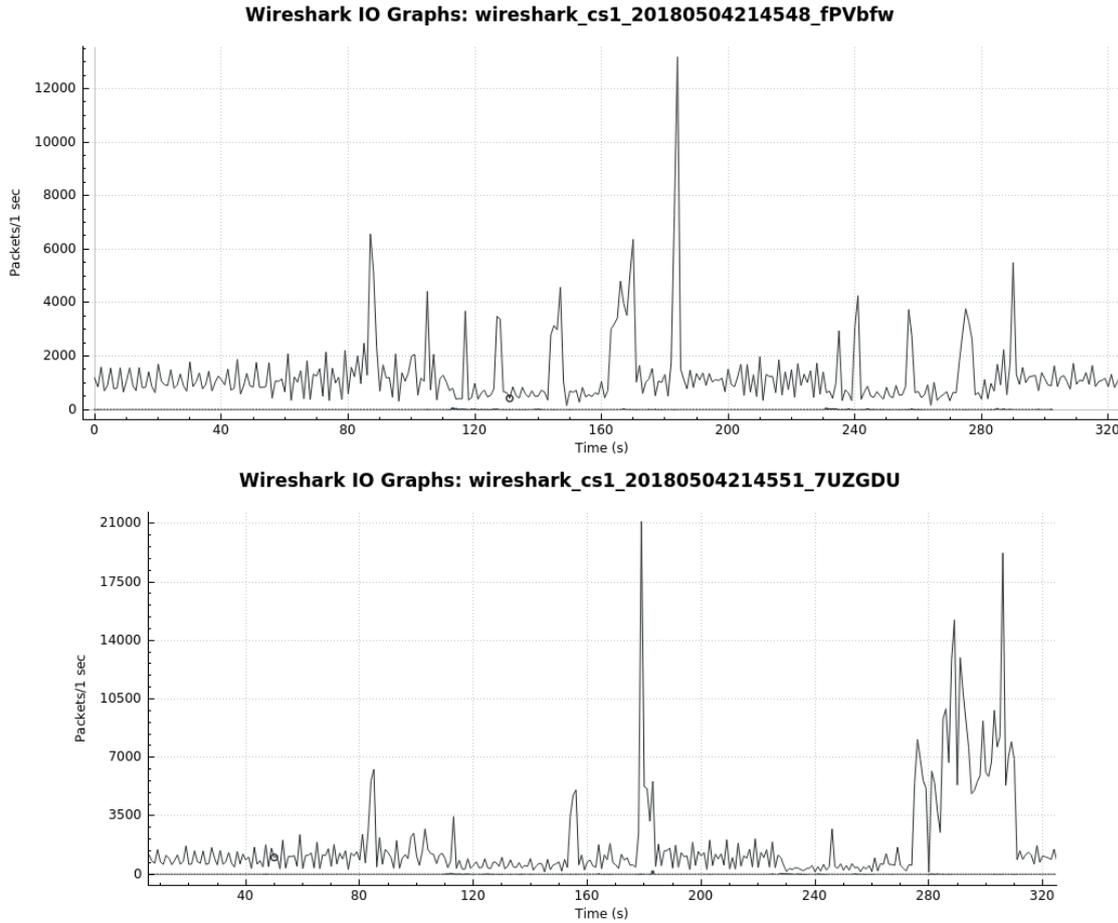


FIGURE 3.7 – Scénario 1 : Trafic inter-contrôleur capturé (en paquets par seconde) pendant la période du scénario de test (à l’aide de Wireshark)

important dans le cas de l’ONOS-WAC. Cette augmentation du trafic est due au processus d’anti-entropie qui a été déclenché par une valeur d’incohérence qui a dépassé le seuil de l’application. A l’inverse, dans le réseau ONOS, le trafic d’anti-entropie est généré périodiquement pendant la période de test, indépendamment des incohérences observées. De même, à $t = 280s$, nous répétons le même scénario en suivant la même séquence d’événements, mais en considérant la défaillance d’un contrôleur différent.

Afin d’évaluer le gain en terme de surcharge du réseau, nous avons estimé le taux ($R(S_i)$) d’augmentation de la surcharge anti-entropie d’ONOS par rapport à ONOS-WAC (voir Equation (3.16)) en fonction du nombre de contrôleurs dans le cluster

$$R_i(S_i) = 1 - \left[\frac{A(S_i) - B(S_i)}{C(S_i) - B(S_i)} \right] \quad (3.16)$$

- $A(S_i)$: surcharge inter-contrôleur générée par ONOS-WAC *après* la séquence d’événements suivant le scénario S_i .

$B(S_i)$: le trafic inter-contrôleurs générés par ONOS-WAC *avant* la séquence d’événements (le trafic inter-contrôleurs global sans le trafic anti-entropie).

$C(S_i)$: surcharge inter-contrôleur générée par ONOS *après* la séquence d’événements suivant le scénario S_i .

Comme le montre la figure 3.8, le gain en surcharge anti-entropie, lors de l’adoption d’ONOS-WAC, croît presque linéairement avec le nombre de contrôleurs dans le cluster avec un gain pouvant atteindre 25%.

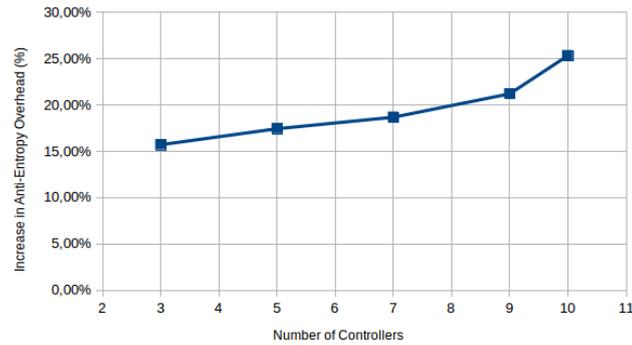


FIGURE 3.8 – Gain en surcharge anti-entropie de l’ONOS-WAC par rapport à l’ONOS par rapport à ONOS en fonction du nombre de contrôleurs dans le cluster

3.6 contribution 3 : modèle de cohérence adaptatif

Dans la continuité de la contribution précédente, nous avons proposé une nouvelle stratégie de réplication de connaissances entre contrôleurs. Cette stratégie consiste en un modèle de cohérence adaptatif basé sur la cohérence éventuelle et qui soit particulièrement adapté aux réseaux à grande échelle. Nous avons ciblé notamment la classe d’applications qui tolèrent des formes et des degrés de cohérence éventuelle assez faible pour des raisons d’évolutivité et de performances, mais qui peuvent néanmoins bénéficier de fonctions de cohérence améliorées. Plus précisément, nous avons proposé une stratégie de réplication évolutive et intelligente suivant les modèles de cohérence répliqués par Quorum. En effet, La stratégie de cohérence peut être modélisée par le mécanisme de réplication basé sur le Quorum. Soit N le nombre total de contrôleurs, N_R la valeur du Quorum de lecture et N_W la valeur du Quorum d’écriture, à chaque fois qu’une mise à jour de l’état du réseau se produit dans un domaine, le contrôleur la transmet à N_W autres contrôleurs. De même, chaque fois qu’un contrôleur effectue une opération de lecture, il extrait les informations sur l’état du réseau de N_R autres contrôleurs. Ainsi, il est possible de mettre en oeuvre deux types de garanties de cohérence.

- Une cohérence forte ou *quorums stricts* exige que la somme du Quorum d’écriture et du Quorum de lecture soit supérieure au nombre total de contrôleurs : $N_W + N_R > N$. Ce modèle de cohérence garantit que l’opération de lecture obtient toujours la dernière valeur de l’élément de données. Cependant, il est coûteux en termes de performances réseau en raison des frais généraux de synchronisation élevés.
- Une cohérence éventuelle ou *quorums partiels* garantie que la somme du Quorum d’écriture et du Quorum de lecture n’ait pas besoin de dépasser le nombre total de répliques : $N_W + N_R \leq N$.

Les systèmes basés sur le Quorum partiel ne garantissent pas que les informations réseau soient le plus à jour possible, ce qui entraîne d’éventuelles informations incohérentes renvoyées aux applications. Cette incohérence peut dépasser le seuil de tolérance admis par certaines applications. Par conséquent, nous appliquons le modèle Probabilistically bounded staleness (PBS) [Bailis et collab., 2012] pour analyser les données de réseau, évaluer le niveau de cohérence, et contrôler l’équilibre entre le niveau de cohérence et les paramètres de réseau (QoS). Sur la base de ces notions, nous avons proposé un modèle de cohérence adaptatif, basé sur le Quorum et utilisant des Quorums partiels. Le modèle proposé utilise les paramètres de réplication comme outil de contrôle, permettant un réglage adaptatif et un contrôle sur les compromis cohérence-performance. Ces compromis en temps réel doivent réduire la surcharge entre les contrôleurs du réseau tout en satisfaisant les exigences de performance et de cohérence de l’application spécifiées dans le SLA donné. L’objectif principal est de trouver, au moment de l’exécution, des configurations optimales de Quorum partiel qui permettent, en fonction de l’évolution des conditions

du réseau et de la charge de l'application, de trouver un compromis entre les exigences continues de l'application en termes de performance (*latence*) et de cohérence (*Obsolescence*). Nous caractérisons l'obsolescence par une métrique d'actualité "Age of Information (AoI)" [Sivasubramanian, 2012] qui décrit la différence entre le temps de requête d'un élément et le temps de sa dernière mise à jour. Si la dernière mise à jour reçue avec succès a été générée au moment $u(t)$, son âge au moment t est $\Delta(t) = t - u(t)$. En outre, nous considérons la latence de la requête de lecture comme notre mesure de performance continue, et nous évaluons également la surcharge (overhead) inter-contrôleur de l'application. Ces compromis en temps réel devraient permettre de minimiser la surcharge totale entre les contrôleurs tout en respectant les seuils définis par l'application et spécifiés dans le contrat de niveau de service donné.

3.6.1 Approche proposée

Afin de mettre en place notre approche adaptative, nous avons proposé un algorithme, basé sur l'apprentissage par renforcement (RL) pour optimiser les deux paramètres N_R et N_W , afin de garantir à la fois les performances du réseau et le niveau de cohérence. L'approche RL utilise l'algorithme Q-Learning [Watkins et Dayan, 1992] (Figure 3.9) qui est un algorithme pour des processus de décision de Markov (MDP)[Sutton et Barto, 2018]. Sur la base de l'expérience et des récompenses, un agent RL peut apprendre des stratégies de contrôle. Le modèle MDP est illustré par un tuple $(\mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R})$ avec \mathcal{S} et \mathcal{A} sont des ensembles d'états et d'actions, $\mathcal{P}(s'|s, a)$ est la probabilité de passer à l'état s' après avoir effectué l'action a à l'état s . La récompense obtenue est $\mathcal{R}(s, a, s')$ après avoir exécuté a à l'état s et être passé à s' . À l'état s , l'action de sélection a est qualifiée par la valeur $Q(s, a)$.

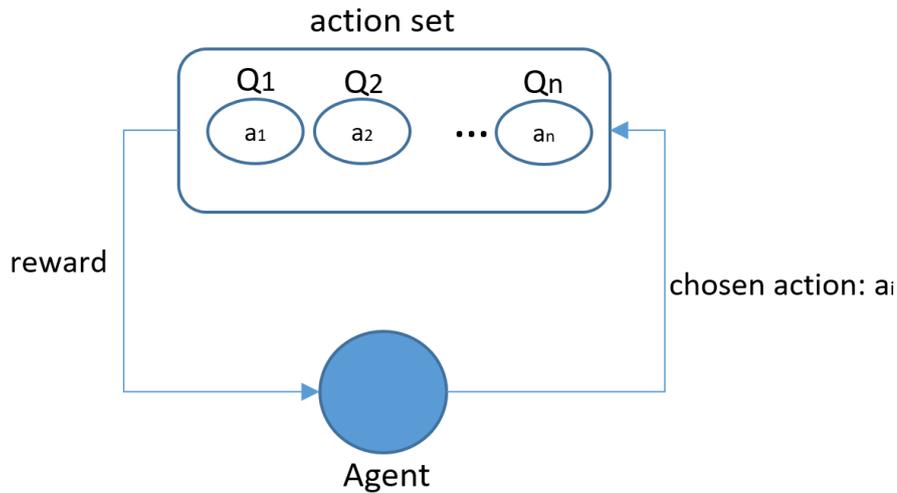


FIGURE 3.9 – Algorithme Q-Learning.

Afin de résoudre un MDP, une politique optimale $\pi : \mathcal{S} \rightarrow \mathcal{A}$ est nécessaire. La politique π détermine l'action correspondant à l'état actuel afin de maximiser les récompenses cumulées. Q (équation de Bellman) peut être calculée comme suit :

$$Q^*(s, a) = r(s, a) + \gamma \sum_{s'} P(s'|s, a) \max_{a'} Q^*(s', a') \quad (3.17)$$

où $r(s, a)$ représente la récompense instantanée attendue de l'action choisie a . γ représente le facteur d'actualisation.

L'algorithme Q-Learning Watkins et Dayan [1992] est proposé pour estimer la fonction action-valeur optimale dans l'équation (3.17). L'agent utilise la récompense retournée r pour mettre à jour les valeurs Q avec l'équation de mise à jour suivante :

$$Q(s, a) = (1 - \alpha)Q(s, a) + \alpha[r + \gamma \max_{a'} Q(s', a')] \quad (3.18)$$

où α est le taux d'apprentissage.

L'approche Q-Learning qui consiste à apprendre une politique qui maximise la fonction action-valeur optimale sans connaissances préalables nous incite à la considérer pour modéliser le problème de sélection de la paire de valeurs N_R et N_W .

Notre approche s'inspire de l'approche TCP Vegas (le mécanisme d'évitement de congestion du TCP [Brakmo et collab., 1994]). Cela s'explique par le fait que l'idée même de TCP Vegas est de contrôler dynamiquement la taille de la fenêtre en fonction du niveau de congestion du réseau. Pour notre problème, les valeurs de N_R et N_W sont ajustées automatiquement en fonction du niveau de cohérence. Si le niveau de cohérence a une tendance à la baisse, les valeurs de N_R et N_W seront augmentées. Dans le cas contraire, si le niveau de cohérence est à la hausse ou acceptable, les valeurs de N_R et N_W seront diminuées.

La récompense est calculée par l'équation (3.19).

$$\text{Récompense}_t = \begin{cases} ThV_{st} - V_{st} & \text{if } WD < ThW \& RD < ThR \\ -ThV_{st} & \text{if } WD > ThW \text{ or } RD > ThR \end{cases} \quad (3.19)$$

où WD et RD sont les métriques du délai de lecture et du délai d'écriture. Le premier est le délai entre le moment où la demande est envoyée à N_R contrôleurs et le moment où les résultats de ces contrôleurs sont renvoyés. Le second est la période de diffusion des messages de mise à jour aux contrôleurs N_W . Les ThW et ThR sont les seuils de délai pour les opérations d'écriture et de lecture. Ils sont considérés comme des seuils tolérables de la latence pour les opérations d'écriture et de lecture.

Nous définissons la variable DIFF pour l'ajustement de N_R et N_W . DIFF est calculé par l'équation (3.20).

$$\text{DIFF} = N * \left(1 - \frac{\text{CL}}{\text{CLbase}} \right) \quad (3.20)$$

La valeur CL représente le niveau de cohérence qui est calculée par l'équation (3.21).

$$\text{CL} = \frac{\pi}{\Pi} \quad (3.21)$$

où CLbase est la plus grande valeur des valeurs CL observées au cours des 10 derniers évènements, π est le nombre de fois où une réplique a lu la dernière valeur au cours des 10 derniers évènements et Π est le nombre total de demandes au cours des 10 derniers évènements.

Les N_R et N_W sont mis à jour avec les équations (3.22) et (3.23). Les deux seuils α et β sont fixés au préalable d'une manière empirique.

$$N_R(t + \delta t) = \begin{cases} N_R(t) - 1 & \text{if } \text{DIFF} < \alpha \\ N_R(t) & \text{if } \alpha < \text{DIFF} < \beta \\ N_R(t) + 1 & \text{if } \text{DIFF} > \beta \end{cases} \quad (3.22)$$

$$N_W(t + \delta t) = \begin{cases} N_W(t) - 1 & \text{if } \text{DIFF} < \alpha \\ N_W(t) & \text{if } \alpha < \text{DIFF} < \beta \\ N_W(t) + 1 & \text{if } \text{DIFF} > \beta \end{cases} \quad (3.23)$$

Pour le mécanisme de sélection de l'algorithme Q-Learning, nous avons appliqué la politique ϵ -greedy [Gomes et Kowalczyk, 2009]. Cette dernière choisit avec une probabilité de $(1 - \epsilon)$ la valeur de N_W ou N_R avec la plus grande valeur Q et avec une probabilité de ϵ une valeur aléatoire de N_W ou N_R .

L'algorithme proposé devrait garantir des valeurs acceptables d'obsolescence de version et d'obsolescence de temps tout en maintenant les métriques WD et RD en dessous des deux seuils ThW et ThD .

3.6.2 Validation expérimentale

Pour évaluer l'efficacité du modèle de cohérence adaptatif proposé nous avons choisi une topologie à grande échelle, appelée *Cogent Topology Zoo*. Cette topologie, représente un réseau situé aux États-Unis et en Europe et qui se compose de 176 commutateurs et de 21 hôtes. Nous avons utilisé l'outil mininet pour émuler la topologie et la diviser en 18 domaines SDN séparés, chacun étant géré par un contrôleur dédié (ONOS, OpenDaylight ou Faucet). En ce qui concerne les caractéristiques des liens, le débit, la latence et la perte sont fixés respectivement à 10 Mb/s, 10 ms et 2 %.

L'outil open-source **Klassen** est utilisé pour générer les flux de données. Cet outil est généralement utilisé pour rejouer un trafic réseau précédemment capturé. Il nous a aidé à générer quelques flux de données initiaux et à les rejouer plusieurs fois par la suite en garantissant les mêmes conditions de comparaison.

Afin d'émuler l'environnement dynamique du réseau, les modifications constantes de la topologie du réseau sont basées sur le modèle classique de Markov à deux états [Haßlinger et Hohlfeld, 2008]. Le premier représente l'état dans lequel le commutateur/liens fonctionne normalement. Le second représente l'état où l'interrupteur/le lien est cassé. La transition d'état est représentée sur la figure 3.10.

Les valeurs de la probabilité de transition p et r sont déterminées par les équations (3.24) et (3.25).

$$p = P(q_t = \text{OFF} | q_{t-1} = \text{ON}) \quad (3.24)$$

$$r = P(q_t = \text{ON} | q_{t-1} = \text{OFF}) \quad (3.25)$$

où q_t est l'état au temps t .

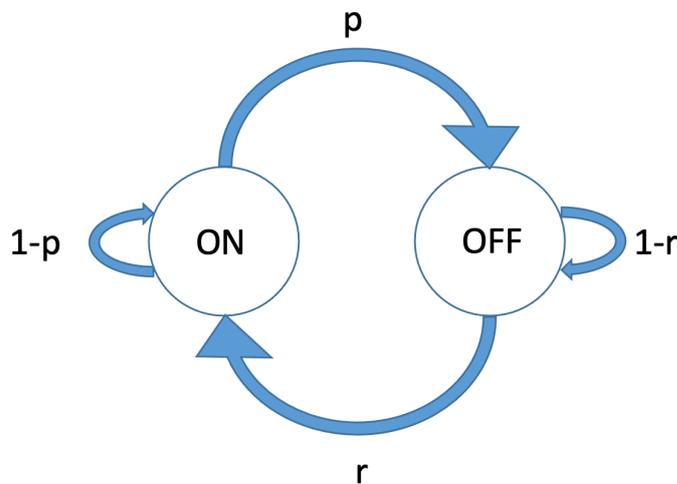


FIGURE 3.10 – Transition d'état des commutateurs et des liens

D'après les expériences, les deux valeurs p et r sont fixées à 50%.

Pour les métriques de performance, elle sont définis suit :

- Métriques liées à la QoS : En ce qui concerne les métriques liées à la QoS, nous nous basons sur l'overhead et les latences. Concrètement, il s'agit des métriques *WriteOverhead*, *ReadOverhead*, *WriteDelay*, et *ReadDelay*. Le *WriteOverhead* et le *ReadOverhead* sont les nombres de paquets inter-contrôleurs capturés sur le port TCP 8787 pour les commandes de lecture et d'écriture. Le *WriteDelay* est la latence du système pour diffuser le message de mise à jour de la topologie à N_W contrôleurs. Le *ReadDelay* est la latence pour envoyer et recevoir des requêtes de lecture depuis N_R .
- Métriques liées à la cohérence : Deux métriques d'obsolescence sont utilisées : l'obsolescence basée sur le temps ($t_staleness$) et l'obsolescence basée sur la version ($v_staleness$)

[Bailis et collab., 2012, 2014]. La première est calculée par la période entre la dernière mise à jour de la base de données locale et le moment où le contrôleur effectue une opération de lecture dessus. La seconde est déterminée par la différence de version entre la version réelle de la base de données locale et celle du moment où le contrôleur effectue une opération de lecture.

Pour la première étape, les deux paramètres N_W et N_R sont fixés aux valeurs de 18 et 1. En d'autres termes, nous réévaluons notre proposition avec la méthode de réplication active en satisfaisant la condition de cohérence forte : $N_W + N_R > N$ (le nombre de contrôleurs dans notre expérience est de 18). Les résultats obtenus pour *WriteDelay* et *ReadDelay* sont illustrés dans la figure 3.11. Le résultat montre que la courbe de *ReadDelay* est toujours stable avec une faible valeur de 2 (ms) car N_R est égal 1. En d'autres termes, cette période courte et stable s'explique par le fait que chaque contrôleur demande les données à un seul et unique contrôleur. La valeur du *WriteDelay* diminue progressivement de 95 à 70 (ms). Sa stabilité se reflète également dans la diminution de l'intervalle de confiance de 7 à 1,9 (ms).

Dans les expériences suivantes, le mécanisme Q-Learning est testé avec les paramètres $\gamma = 0.8$ et $\alpha = 0.6$. Au départ, les valeurs initiales de N_R et N_W sont respectivement de 4 et 14. La valeur du seuil d'écriture est fixée à 50 (ms). Trois scénarios sont lancés avec trois valeurs différentes du seuil de lecture : 10 ms, 30 ms et 50 ms. La valeur ϵ du mécanisme de sélection $\epsilon - greedy$ est fixée à 10%.

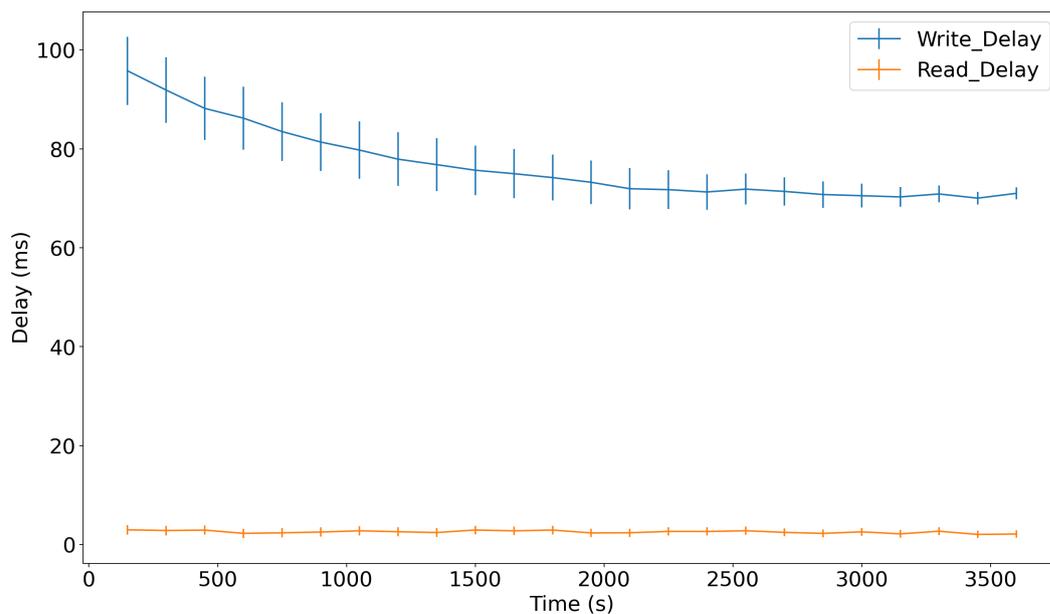


FIGURE 3.11 – ReadDelay et WriteDelay avec les valeurs fixes de $N_W = 18$ et $N_R = 1$.

Comme le montre la figure 3.12, bien que la sélection adaptative de la valeur N_R augmente le délai de lecture par rapport à la méthode de réplication active, ces valeurs de délai ne dépassent pas les seuils fixés à 10 ms, 20 ms ou 30 ms. Après environ 1200s d'initialisation, la variation du délai liée aux trois choix de seuils effectués devient stable. Par ailleurs, la sélection adaptative de la valeur N_W de l'algorithme proposé diminue le délai d'écriture. En conséquence, le délai d'écriture dans les trois cas est toujours inférieur à la valeur du seuil de 50 ms. Concrètement, les valeurs de convergence de la valeur moyenne du délai d'écriture pour les trois cas sont, respectivement, de 45,2s, 40s et 38,3 s. Ce résultat montre également que plus le seuil de lecture est augmenté, plus la valeur du délai d'écriture est réduite.

L'obsolescence de version et l'obsolescence de temps sont illustrées dans la Figure 3.13. Après 1500s, la valeur de la version devient stable. On constate que la métrique de l'obsolescence de version est toujours inférieure au seuil fixe (0,5). La courbe correspondant au seuil de lecture de

30 ms atteint la plus petite valeur de l'intervalle, ce qui montre l'importance de la sélection de la valeur du seuil de lecture. Une sélection appropriée du seuil de lecture donne un bon niveau de cohérence en termes de stabilité des versions. En ce qui concerne les résultats obtenus pour l'obsolescence de temps, après les 1500ms de fluctuation, on constate que la valeur de cette dernière revient à un état stable.

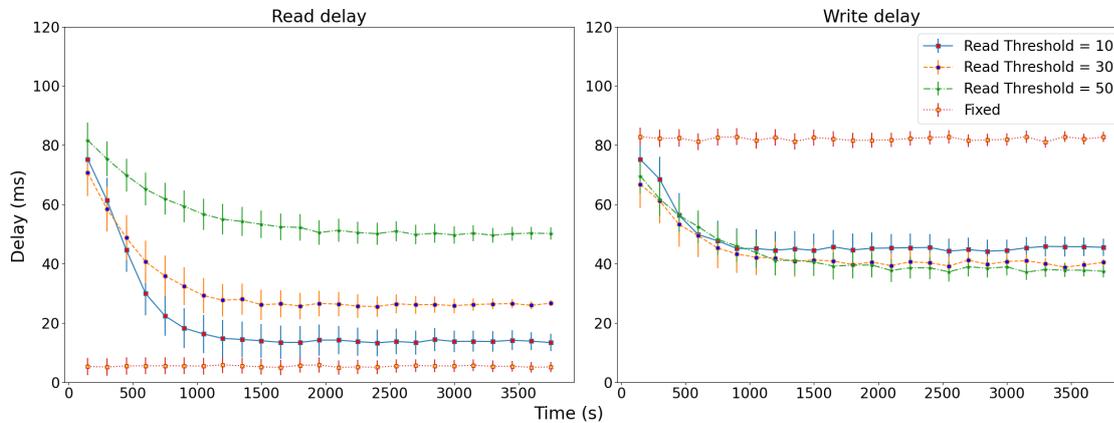


FIGURE 3.12 – ReadDelay et WriteDelay avec différentes valeurs de ThR

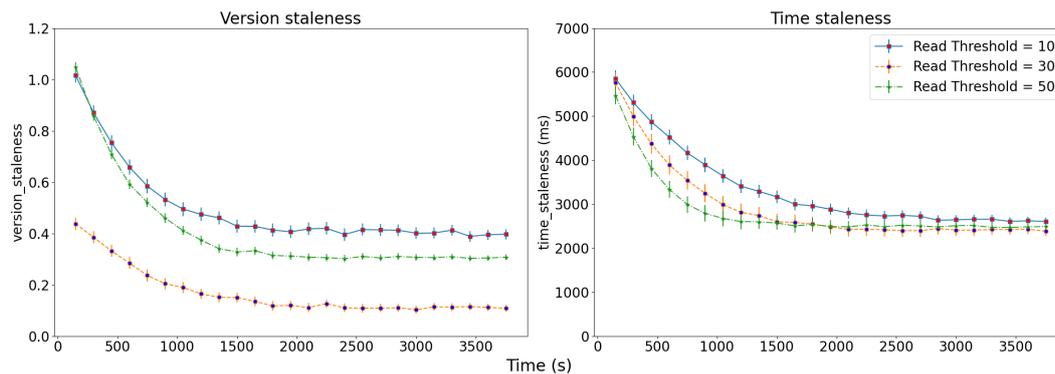
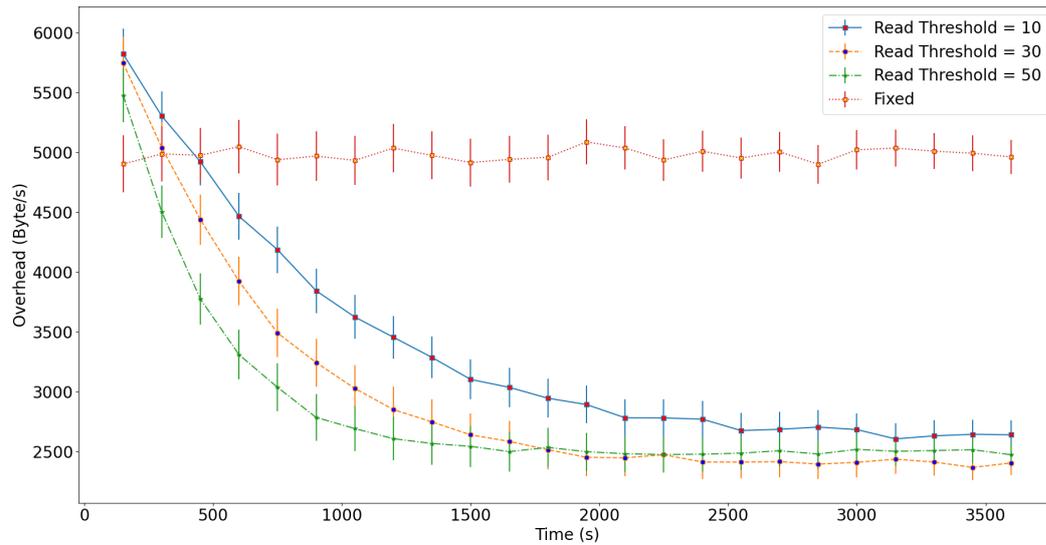


FIGURE 3.13 – Version staleness et time staleness avec différentes valeurs de ThR

La figure 3.14 montre les résultats en terme d'overhead. On constate que le cas fixe obtient toujours la valeur d'overhead la plus élevée de 4950 (octets/s) car la somme de ses deux valeurs N_W et N_R est supérieure aux autres. Après environ 1500 s, les trois courbes correspondant aux seuils de lecture de 10 ms, 30 ms et 50 ms deviennent progressivement stables avec les valeurs moyennes convergentes de 2670, 2415 et 2500 (octets/s).

Ces résultats expérimentaux montrent que la solution proposée pouvait maintenir un niveau de cohérence acceptable, en terme d'obsolescence de version et de temps, tout en contrôlant continuellement la limite des différentes mesures du réseau telles que le délai de lecture, le délai d'écriture et l'overhead.

FIGURE 3.14 – Survol avec différentes valeurs de ThR

3.7 Conclusion et travail réalisé

Le réseau défini par logiciel (SDN) a de plus en plus de succès dans le milieu académique ainsi que dans la recherche. Le paradigme SDN fonde ses promesses sur la séparation des préoccupations entre la logique de contrôle du réseau et les équipements de transmission réseau, ainsi que sur la centralisation logique de l'intelligence réseau dans des composants logiciels. Grâce à ces attributs clés, on estime que l'approche SDN, avec notamment les nouvelles technologies de virtualisation réseau (NFV), permettra de changer fondamentalement le paysage réseau et facilitera ainsi le passage à des réseaux de nouvelle génération plus flexibles, agiles, programmables, adaptables et hautement automatisés.

L'approche fondée sur le SDN soulève, néanmoins, de nombreuses préoccupations et questions concernant sa mise en œuvre et son déploiement. Par exemple, les déploiements SDN actuels basés sur des architectures de contrôle physiquement centralisées soulèvent encore plusieurs problèmes d'évolutivité et de fiabilité. Par conséquent, les architectures de contrôle SDN distribuées ont été proposées comme solution appropriée pour tenter de résoudre ces problèmes. Il n'empêche qu'il existe encore des débats communautaires sur la meilleure approche à adopter pour décentraliser le plan de contrôle du réseau afin d'exploiter pleinement le potentiel du SDN. Ces discussions portent principalement sur les différents compromis et challenges impliqués dans la décentralisation du plan de contrôle.

Au démarrage de ces travaux nous avons mené une étude bibliographique sur la vaste gamme de plates-formes de contrôleurs SDN existantes. Outre la révision du concept SDN et l'étude de son architecture en la confrontant à l'architecture réseau traditionnelle, nous avons proposé une taxonomie des plates-formes de contrôleurs SDN, en les catégorisant de deux manières : sur la base d'une classification physique ou d'une classification logique. Notre étude approfondie de ces propositions de plate-formes SDN nous a permis de mieux comprendre leurs avantages et inconvénients et de développer une prise de conscience critique des défis du contrôle distribué dans les SDNs.

En particulier, l'évolutivité, la fiabilité, la cohérence et l'interopérabilité du plan de contrôle SDN figurent parmi les principaux défis à relever pour la conception d'une plate-forme de contrôleurs SDN distribués robuste et performante. Bien que considérées comme les principales limites des conceptions de contrôle SDN totalement centralisées, l'évolutivité et la fiabilité sont également des préoccupations majeures qui sont exprimées dans le contexte des architectures SDN distribuées. Elles sont, en effet, fortement impactées par la structure du plan de contrôle distribué (par exemple, une organisation plate, hiérarchique ou hybride) ainsi que par le nombre et le

placement des multiples contrôleurs au sein du réseau SDN. La réalisation de telles exigences en matière de performance et de disponibilité se fait généralement au détriment de la garantie d'une vue de réseau centralisée cohérente, nécessaire au bon comportement des applications SDN. Il convient donc d'examiner les considérations liées à la cohérence parmi les compromis impliqués dans le processus de conception d'une plate-forme décentralisée de contrôleurs SDN.

Compte tenu de la diversité des plate-formes prometteuses de contrôleurs SDN et de leur vaste éventail de défis majeurs, nous sommes convaincus que le développement d'une nouvelle plate-forme n'est peut-être pas la meilleure solution. Cependant, il est essentiel de tirer parti des plates-formes existantes en agrégeant, fusionnant et améliorant les idées proposées afin de se rapprocher le plus possible d'une norme commune qui pourrait émerger dans les années à venir. Cette plate-forme de contrôleurs SDN distribués devrait répondre aux défis émergents associés aux déploiements à grande échelle et, plus important encore, aux réseaux de prochaine génération (par exemple, IoT [Ojo et collab. \[2016\]](#) et Fog Computing [Liang et collab. \[2017\]](#)).

C'est dans cette optique que nous avons abordé certains des problèmes évoqués précédemment et associés au problème complexe de la conception d'un plan de contrôle SDN distribué. Pour ce faire, nous avons proposé de scinder ce problème en deux parties gérables et corrélées : le placement de contrôleurs (1) et le partage des connaissances (2). La première partie s'occupe d'étudier le nombre requis de contrôleurs ainsi que leurs emplacements appropriés par rapport aux objectifs de performance et de fiabilité souhaités et en fonction des contraintes existantes. La seconde partie est liée au type et à la quantité d'informations à partager entre les instances de contrôleurs en fonction du niveau de cohérence souhaité.

Tout d'abord, nous avons abordé le problème d'optimisation de placement des contrôleurs SDN dans le contexte de réseaux à large échelle de type IoT. Pour ce faire, nous avons étudié plusieurs stratégies qui couvrent différents aspects du problème d'optimisation multi-objectifs de l'emplacement des contrôleurs SDN au regard de multiples métriques de fiabilité et de performance prises en compte en fonction de différents usages et contextes. Pour évaluer ces stratégies, deux approches heuristiques ont été proposées dans le but de trouver des solutions approximatives de haute qualité au problème de placement des contrôleurs dans un temps de calcul raisonnable : une approche de partitionnement (PAM-B) basée sur un score de dissimilarité et une approche génétique modifiée (NSGA-II). Nos résultats ont démontré le potentiel des techniques de partitionnement dans la perspective de fournir des configurations de placement de contrôleurs appropriées qui permettent un compromis équilibré entre les critères concurrents de performance et de fiabilité à grande échelle.

Ensuite, nous avons étudié le problème de partage des connaissances entre les contrôleurs SDN distribués en proposant un modèle de cohérence adaptatif à plusieurs niveaux, basé sur la notion de cohérence continue pour les contrôleurs SDN distribués. Ce modèle présente de nombreux avantages pour les applications SDN par rapport aux extrêmes de cohérence forte et de cohérence éventuelle, en particulier dans les déploiements à large échelle : Il offre les avantages d'évolutivité, de performance et de disponibilité d'un modèle de cohérence éventuel, mais présente l'avantage supplémentaire de contrôler les incohérences d'état observées d'une manière spécifique à l'application. Plus spécifiquement, nous avons proposé deux différentes approches de cohérence évolutives pour les contrôleurs open-source ONOS et nous les comparons avec les stratégies statiques d'ONOS pour la cohérence d'état éventuelle.

La première approche de cohérence a été mise en œuvre pour une application de routage à la source au-dessus d'ONOS. Elle consiste à transformer le modèle de cohérence éventuelle d'ONOS en un modèle de cohérence adaptative en utilisant la *période de réconciliation Anti-Entropie* comme un bouton de contrôle pour un réglage fin et adaptatif des niveaux de cohérence. En plus de garantir les exigences de cohérence continues de l'application (c'est-à-dire les limites d'erreur numérique) spécifiées dans le contrat de niveaux de service d'application donné (SLA), nos résultats ont montré une réduction substantielle de la surcharge due à la réconciliation Anti-Entropie par rapport à l'approche de cohérence statique à l'échelle d'ONOS.

La deuxième approche étend la stratégie de cohérence adaptative à la technique de répllication

optimiste utilisée dans le modèle de cohérence éventuelle d'ONOS. Elle a été implémentée pour une application de type CDN que nous avons développée sur les contrôleurs ONOS. Elle consiste principalement à transformer la technique de réplication optimiste en une stratégie de réplication inspirée du Quorum, plus évolutive et plus intelligente, portant sur diverses approches d'apprentissage par renforcement, plus précisément de Q-learning : elle utilise notamment *les paramètres de Quorum partiel en lecture/écriture* comme *paramètres de contrôle ajustables* pour un réglage fin de la cohérence, plutôt que de reposer sur des mécanismes de réconciliation Anti-Entropie. Nos expérimentations opérées sur 3 contrôleurs (Faucet, ODL et ONOS) ont montré que l'approche proposée s'avérait efficace pour trouver, au moment de l'exécution, les paramètres de réplication Quorum en lecture/écriture appropriés permettant d'obtenir, dans des conditions changeantes de réseau et de charge de travail de l'application, des compromis équilibrés entre les exigences continues de l'application en matière de performance (*latence*) et de cohérence (*Obsolescence*). Ces compromis en temps réel ont permis de réduire considérablement la surcharge totale (entre contrôleurs) liée à l'application, tout en satisfaisant les exigences spécifiées dans les SLAs.

3.8 Synthèses

3.8.1 Difficultés rencontrées

L'une des difficultés les plus importantes rencontrées dans ce travail consiste dans le fait qu'il n'a pas toujours été facile d'expliquer l'intérêt de rendre une architecture distribuée alors même que l'objectif principal des réseaux SDN fut la centralisation du contrôle. Néanmoins, nous avons assisté à une prise de conscience progressive de la part de la communauté, ce qui a grandement facilité l'acceptabilité de nos propositions.

3.8.2 Forces et faiblesses

Selon moi, l'intérêt principal de notre modèle réside dans sa praticité. En effet, l'approche repose sur des outils logiciels standards et utilisés aussi bien en recherche qu'en industrie.

Néanmoins la mise en pratique de ces modèles est assez dépendante des choix technologiques choisis et si on considère l'évolution permanente de ces derniers, la maintenance des implémentations proposés devient difficile.

3.8.3 Possible extension

Les travaux que nous avons menés peuvent être étendus de deux manières : La première consiste à réfléchir à une évolution de l'architecture permettant de rendre la conception agnostique et qu'elle ne soit plus dépendante d'un contrôleur SDN donné.

La seconde piste consisterait à prendre en compte le volet sécurité en assurant une traçabilité et un suivi des connaissances partagés par les contrôleurs.

3.8.4 Impact potentiel des travaux

Les travaux présentés dans ce chapitre ont tous fait l'objet de publications qui furent citées plus de 300 fois par nos pairs. De plus, nos codes sources ont à chaque fois été partagés afin de : 1/ permettre une analyse et un retour sur nos différents implémentations par nos pairs, 2/ donner un cadre de travail à ceux qui souhaitent entreprendre des travaux et contribuer à ces problématiques.

3.8.5 Retour d'expérience

Ayant été confronté au début de nos travaux à des difficultés pour passer de la théorie à la pratique, nous avons noué des liens et des collaborations avec l'open networking foundation, or-

ganisme qui supporte le développement du contrôleur ONOS. Avec le recul, il est évident que nous n'aurions pas pu appliquer de telles approches sans un pareil cadre et qu'il aurait été difficile de faire admettre la pertinence de nos idées sans un ancrage industriel. Nous devons aussi noter l'intérêt de partager non seulement les idées mais également les implémentations et codes sources dans un objectif, de reproductibilité et de transparence mais également afin de faciliter le travail des jeunes collègues qui souhaitent travailler sur ces thématiques.

Bibliographie

- «Pox», <http://www.noxrepo.org/pox/about-pox/>. Accessed : 2015-11-22. 26
- 2015, «Floodlight Project», URL <http://www.projectfloodlight.org/>, accessed : 2015-12-07. 26
- 2016, «Opendaylight project», URL <http://www.opendaylight.org/>, accessed : 2016.-01-05. 26, 27, 35
- ONOS. URL <https://onosproject.org/>. 35
- Ahmadi, V., A. Jalili, S. M. Khorramizadeh et M. Keshtgari. 2015, «A hybrid nsga-ii for solving multiobjective controller placement in sdn», dans *The 2nd International Conference on Knowledge-Based Engineering and Innovation (KBEI)*, p. 663–669, doi :10.1109/KBEI.2015.7436122. 29
- Aslan, M. et A. Matrawy. 2016, «Adaptive consistency for distributed SDN controllers», dans *2016 17th International Telecommunications Network Strategy and Planning Symposium (Networks)*, p. 150–157, doi :10.1109/NETWKS.2016.7751168. 35
- Bailis, P., S. Venkataraman, M. J. Franklin, J. M. Hellerstein et I. Stoica. 2012, «Probabilistically bounded staleness for practical partial quorums», *CoRR*, vol. abs/1204.6082. URL <http://arxiv.org/abs/1204.6082>. 39, 43
- Bailis, P., S. Venkataraman, M. J. Franklin, J. M. Hellerstein et I. Stoica. 2014, «Quantifying eventual consistency with pbs», *The VLDB Journal*, vol. 23, n° 2, p. 279–302. 43
- Bannour, F., S. Souihi et A. Mellouk. 2017, «Scalability and Reliability Aware SDN Controller Placement Strategies», dans *Proc. Of the 13th International Conf. on Network and Service Management, CNSM 2017, MANSDDNFV2017*, Tokyo, Japan. URL <https://hal.archives-ouvertes.fr/hal-01676567>. 30
- Bannour, F., S. Souihi et A. Mellouk. 2018, «Distributed sdn control : Survey, taxonomy, and challenges», *IEEE Communications Surveys Tutorials*, vol. 20, n° 1, doi :10.1109/COMST.2017.2782482, p. 333–354. 26
- Berde, P., M. Gerola, J. Hart, Y. Higuchi, M. Kobayashi, T. Koide, B. Lantz, B. O'Connor, P. Radoslavov, W. Snow et G. Parulkar. 2014, «Onos : Towards an open, distributed SDN OS», dans *Proceedings of the Third Workshop on Hot Topics in Software Defined Networking, HotSDN '14*, ACM, New York, NY, USA, ISBN 978-1-4503-2989-7, p. 1–6, doi :10.1145/2620728.2620744. 26, 27
- Botelho, F., A. Bessani, F. M. V. Ramos et P. Ferreira. 2014, «On the design of practical fault-tolerant sdn controllers», dans *2014 Third European Workshop on Software Defined Networks*, ISSN 2379-0350, p. 73–78, doi :10.1109/EWSDN.2014.25. 26
- Brakmo, L. S., S. W. O'Malley et L. L. Peterson. 1994, «Tcp vegas : New techniques for congestion detection and avoidance», vol. 24, n° 4, doi :10.1145/190809.190317, p. 24–35, ISSN 0146-4833. URL <https://doi.org/10.1145/190809.190317>. 41

- Clark, D. D., C. Partridge, J. C. Ramming et J. T. Wroclawski. 2003, «A knowledge plane for the internet», dans *Proceedings of the 2003 conference on Applications, technologies, architectures, and protocols for computer communications*, p. 3–10. [24](#)
- Curtis, A. R., J. C. Mogul, J. Tourrilhes, P. Yalagandula, P. Sharma et S. Banerjee. 2011, «Devoflow : Scaling flow management for high-performance networks», dans *Proceedings of the ACM SIGCOMM 2011 Conference*, SIGCOMM '11, ACM, New York, NY, USA, ISBN 978-1-4503-0797-0, p. 254–265, doi :10.1145/2018436.2018466. [26](#)
- Das, T., V. Sridharan et M. Gurusamy. 2019, «A survey on controller placement in SDN», *IEEE Communications Surveys Tutorials*, doi :10.1109/COMST.2019.2935453, p. 1–1. [27](#)
- Feamster, N., J. Rexford et E. Zegura. 2014, «The road to sdn : An intellectual history of programmable networks», *SIGCOMM Comput. Commun. Rev.*, vol. 44, n° 2, doi :10.1145/2602204.2602219, p. 87–98, ISSN 0146-4833. URL <https://doi.org/10.1145/2602204.2602219>. [24](#)
- Ganatra, N. S. A. 2012, «Comparative study of several clustering algorithms.», *International Journal of Advanced Computer Research*, p. 37–42. [29](#)
- Gomes, E. R. et R. Kowalczyk. 2009, «Dynamic analysis of multiagent q-learning with epsilon-greedy exploration», dans *Proceedings of the 26th International Conference on Machine Learning*. [41](#)
- Gude, N., T. Koponen, J. Pettit, B. Pfaff, M. Casado, N. McKeown et S. Shenker. 2008, «Nox : Towards an operating system for networks», *SIGCOMM Comput. Commun. Rev.*, vol. 38, n° 3, doi :10.1145/1384609.1384625, p. 105–110, ISSN 0146-4833. [26](#)
- Gupta, A., M. Shahbaz, L. Vanbever, H. Kim, R. Clark, N. Feamster, J. Rexford et S. Shenker. 2014, «Sdx : A software defined internet exchange», *ACM SIGCOMM*. [26](#)
- Hassas Yeganeh, S. et Y. Ganjali. 2012, «Kandoo : A framework for efficient and scalable offloading of control applications», dans *Proceedings of the First Workshop on Hot Topics in Software Defined Networks*, HotSDN '12, ACM, New York, NY, USA, ISBN 978-1-4503-1477-0, p. 19–24, doi :10.1145/2342441.2342446. [26](#)
- Haßlinger, G. et O. Hohlfeld. 2008, «The gilbert-elliott model for packet loss in real time services on the internet», dans *14th GI/ITG Conference-Measurement, Modelling and Evaluation of Computer and Communication Systems*, VDE, p. 1–15. [42](#)
- Heller, B., R. Sherwood et N. McKeown. 2012, «The controller placement problem», dans *Proceedings of the First Workshop on Hot Topics in Software Defined Networks*, HotSDN '12, ACM, New York, NY, USA, ISBN 978-1-4503-1477-0, p. 7–12, doi :10.1145/2342441.2342444. [28](#)
- Hock, D., M. Hartmann, S. Gebert, M. Jarschel, T. Zinner et P. Tran-Gia. 2013, «Pareto-optimal resilient controller placement in sdn-based core networks», dans *25th International Teletraffic Congress (ITC)*, Shanghai, China. [30](#)
- Jain, S., A. Kumar, S. Mandal, J. Ong, L. Poutievski, A. Singh, S. Venkata, J. Wanderer, J. Zhou, M. Zhu, J. Zolla, U. Hölzle, S. Stuart et A. Vahdat. 2013, «B4 : Experience with a globally-deployed software defined wan», dans *Proceedings of the ACM SIGCOMM 2013 Conference on SIGCOMM*, SIGCOMM '13, ACM, New York, NY, USA, ISBN 978-1-4503-2056-6, p. 3–14, doi : 10.1145/2486001.2486019. [26](#)
- Katta, N., H. Zhang, M. Freedman et J. Rexford. 2015, «Ravana : Controller fault-tolerance in software-defined networking», dans *Proceedings of the 1st ACM SIGCOMM Symposium on Software Defined Networking Research*, SOSR '15, ISBN 978-1-4503-3451-8, p. 4 :1–4 :12. [26](#)

- Killi, B. P. R. et S. V. Rao. 2019, «Controller placement in software defined networks : A comprehensive survey», *Computer Networks*, vol. 163, p. 106-883, ISSN 1389-1286. 28
- Klassen, F. *TcpReplay*. <https://tcpreplay.appneta.com>. 42
- Koponen, T., M. Casado, N. Gude, J. Stribling, L. Poutievski, M. Z. Google, R. Ramanathan, Y. I. NEC, H. I. NEC, T. H. NEC et S. Shenker. 2010, «Onix : a distributed control platform for large-scale production networks», dans *9th Conference on Operating Systems Design and Implementation*, p. 351–364. 26, 27
- Lakshman, A. et P. Malik. 2010, «Cassandra : A decentralized structured storage system», *SIGOPS Oper. Syst. Rev.*, vol. 44, n° 2, doi :10.1145/1773912.1773922, p. 35–40, ISSN 0163-5980. 35
- Lange, S., S. Gebert, T. Zinner, P. Tran-Gia, D. Hock, M. Jarschel et M. Hoffmann. 2015, «Heuristic approaches to the controller placement problem in large scale sdn networks.», *IEEE Transactions on Network and Service Management*, vol. 12, n° 1, p. 4–17. 28
- Liang, K., L. Zhao, X. Chu et H. H. Chen. 2017, «An integrated architecture for software defined and virtualized radio access networks with fog computing», *IEEE Network*, vol. 31, n° 1, doi : 10.1109/MNET.2017.1600027NM, p. 80–87, ISSN 0890-8044. 46
- Ojo, M., D. Adami et S. Giordano. 2016, «A SDN-IoT architecture with NFV implementation», dans *2016 IEEE Globecom Workshops (GC Wkshps)*, p. 1–6, doi :10.1109/GLOCOMW.2016.7848825. 46
- ONF. «Open Networking Foundation», <https://www.opennetworking.org/>. Accessed : 2016-05-19. 24
- Ongaro, D. et J. Ousterhout. 2014, «In search of an understandable consensus algorithm», dans *2014 USENIX Annual Technical Conference (USENIX ATC 14)*. 35
- Phemius, K., M. Bouet et J. Leguay. 2013, «DISCO : distributed multi-domain SDN controllers», *CoRR*, vol. abs/1308.6138. 26
- Shraddha, K. et M. Emmanuel. 2014, «Review and comparative study of clustering techniques.», *International Journal of Computer Science and Information Technology (IJCSIT)*, vol. 5, p. 805–812. 33
- Sivasubramanian, S. 2012, «Amazon dynamodb : A seamlessly scalable non-relational database service», dans *Proceedings of the 2012 ACM SIGMOD International Conference on Management of Data*, SIGMOD '12, ACM, New York, NY, USA, ISBN 978-1-4503-1247-9, p. 729–730, doi :10.1145/2213836.2213945. 35, 40
- Sutton, R. S. et A. G. Barto. 2018, *Reinforcement Learning : An Introduction*, 2^e éd., The MIT Press. URL <http://incompleteideas.net/book/the-book-2nd.html>. 40
- Tootoonchian, A. et Y. Ganjali. 2010, «Hyperflow : A distributed control plane for openflow», dans *Proceedings of the 2010 Internet Network Management Conference on Research on Enterprise Networking*, INM/WREN'10, Berkeley, CA, USA, p. 3–3. 26, 27
- Watkins, C. J. C. H. et P. Dayan. 1992, «Q-learning», dans *Machine Learning*, p. 279–292. 40
- Yu, H. et A. Vahdat. 2000, «Design and evaluation of a continuous consistency model for replicated services», dans *Proceedings of the 4th Conference on Symposium on Operating System Design & Implementation - Volume 4*, OSDI'00, Berkeley, CA, USA. 35, 36
- Yu, M., J. Rexford, M. J. Freedman et J. Wang. 2010, «Scalable flow-based networking with difane», dans *Proceedings of the ACM SIGCOMM 2010 Conference*, SIGCOMM '10, ACM, New York, NY, USA, ISBN 978-1-4503-0201-2, p. 351–362, doi :10.1145/1851182.1851224. 26

Chapitre 4

Supervision du réseau : de la collecte à l'élaboration des connaissances

« *Power is gained by sharing knowledge, not hoarding it.* »

Unknown

Resumé

Dans ce chapitre, **on s'intéresse aux connaissances nécessaires aux opérateurs afin de contrôler et gérer les infrastructures réseaux**. On se concentre principalement sur 2 aspects : 1/ l'évolution naturelle du trafic qui tend de plus en plus à être chiffrée et 2/ l'estimation de la qualité telle qu'elle est perçue par l'utilisateur qui tend à s'imposer comme la métrique phare du point de vue des opérateurs.

Sommaire

4.1 Introduction et contexte	52
4.2 Problématique : Collecte des données et élaboration des connaissances	52
4.3 Contribution 1 : Analyse et classification du trafic	52
4.3.1 État de l'art de la classification du trafic	52
4.3.2 Une nouvelle méthode de classification du trafic basée sur le réseau de neurones convolutifs	53
4.3.3 Validation expérimentale	56
4.4 Contribution 2 : Estimation de la QoE	61
4.4.1 Estimation de la QoE via des techniques de ML	61
4.4.2 Nouvelle approche d'estimation de la QoE	64
4.4.3 Évaluation du modèle de prédiction BFC	67
4.5 Conclusion et travail réalisé	68
4.6 Synthèses	69
4.6.1 Difficultés rencontrées	69
4.6.2 Forces et faiblesses	69
4.6.3 Possible extension	69
4.6.4 Impact potentiel des travaux	69
4.6.5 Retour d'expérience	69

4.1 Introduction et contexte

Globalement, il est devenu particulièrement difficile pour les réseaux actuels de fournir le niveau requis de qualité de service (QoS), et encore moins la qualité d'expérience (QoE) qui introduit de nouvelles exigences centrées sur l'utilisateur. Pour être plus précis, se fier uniquement à la QoS traditionnelle basée sur des paramètres de performances techniques (e.g. bande passante et latence) s'avère insuffisant pour les réseaux avancés et en expansion d'aujourd'hui. De plus, répondre à ce nombre croissant d'indicateurs de performance est une tâche d'optimisation complexe qui peut être traitée comme un problème NP-complet. Par ailleurs, les opérateurs de réseau réalisent de plus en plus l'importance de l'expérience globale de l'utilisateur final et de sa perception subjective des services fournis qui permettent de résoudre des problèmes que les mécanismes fondés sur la qualité de service ont du mal à résoudre. En conséquence, les tendances actuelles en matière de gestion de réseau se dirigent vers ce nouveau concept, couramment appelé QoE, qui représente la qualité globale d'un service de réseau du point de vue de l'utilisateur final. Néanmoins, compte tenu de l'augmentation du trafic chiffré dans l'Internet, il est devenu assez difficile pour les opérateurs de concevoir des techniques d'estimation de la QoE sans connaissance de la nature du trafic. Pour y répondre, il est donc nécessaire de concevoir des mécanismes intelligents capables dans un premier temps d'identifier la nature du trafic en toutes circonstances afin d'appliquer dans un deuxième temps le mécanisme le plus approprié d'estimation de la QoE.

4.2 Problématique : Collecte des données et élaboration des connaissances

4.3 Contribution 1 : Analyse et classification du trafic

La classification du trafic consiste à identifier les classes d'applications dans le réseau. Dans nos travaux, elle nous permet, par exemple, de mettre en œuvre des mécanismes différenciés (un routage adapté aux applications, des politiques de facturation évoluées, ...). Ces mécanismes sont pilotés par une mesure de Qualité d'expérience qui dépend de la nature de l'application. Traditionnellement, les méthodes de classification de trafic reposaient sur l'identification des ports et sur l'analyse des contenus des paquets. L'identification des ports est une méthode dont l'efficacité est limitée car les applications et les protocoles modernes ne sont pas toujours liés à un port spécifique. En outre, de nombreux réseaux modifient le numéro de port dans les paquets à l'aide de la traduction d'adresses réseau (NAT) et de la traduction de port (PAT). L'analyse des contenus a pour objectif d'identifier des signatures correspondant à une application particulière [Adami et collab., 2015]. Cette approche permettait d'atteindre une grande précision avant que les protocoles de transports n'adoptent une politique de chiffrement systématique. Récemment, Google a développé Quick UDP Internet Connection (QUIC) [Langley et collab., 2017], un nouveau protocole de réseau de couche de transport basé sur UDP. QUIC présente de nombreux avantages liés à l'établissement de la connexion, au contrôle de la congestion, au multiplexage et à la migration de la connexion. En outre, QUIC fournit également une protection de sécurité équivalente à TLS (Transport Layer Security). L'émergence de QUIC a rendu les méthodes traditionnelles de classification du trafic, notamment les méthodes basées sur l'identification des ports et des données utiles, inadaptées.

4.3.1 État de l'art de la classification du trafic

La classification du trafic [Lopez-Martin et collab., 2017] joue un rôle essentiel dans le dépannage du réseau pour les opérateurs de réseau. L'objectif de la classification du trafic est d'identifier les classes d'applications dans le réseau afin de mettre en œuvre des mécanismes adaptés pour

résoudre les problèmes du réseau et répondre aux exigences strictes en termes de SLA. Concrètement, la classification du trafic est largement utilisée dans le routage adapté aux applications, les politiques de facturation, les systèmes de détection des intrusions, l'identification des menaces de sécurité, l'application des exigences de qualité de service, etc. Par conséquent, la classification du trafic a été largement débattu par la communauté des chercheurs.

Dans le passé, de nombreuses tentatives d'identification de la classe d'application reposaient sur des approches basées sur les ports et sur les charges utiles. La première approche utilise les numéros de port dans les paquets TCP/UDP pour identifier la classe d'application. Cependant, elle n'est pas efficace car les applications et les protocoles modernes ne sont pas toujours liés à un port spécifique. En outre, de nombreux réseaux modifient le numéro de port dans les paquets à l'aide des mécanismes de NAT et de PAT, ce qui entraîne des résultats inexacts pour la classification du trafic. La deuxième approche [Adami et collab., 2015] inspecte une charge utile du paquet pour identifier ses signatures correspondant à une application particulière, appelée Deep Packet Inspection. Cette approche peut atteindre une grande précision, mais elle est limitée en raison du trafic réseau crypté. Concrètement, le trafic réseau est aujourd'hui crypté pour protéger les données et la vie privée des utilisateurs, de sorte que la signature de l'application dans la charge utile du paquet est cachée.

Pour surmonter les inconvénients de ces approches, une approche statistique utilisant des outils d'apprentissages automatiques [Lopez-Martin et collab., 2017] est apparue. Cette approche analyse les informations de haut niveau obtenues à partir des flux réseau (par exemple, la taille de la fenêtre TCP, les temps d'inter-arrivée des paquets, etc.) en utilisant le ML pour identifier la classe d'application. Les caractéristiques basées sur les flux peuvent être extraites de tous les paquets d'un flux [Fan et Liu, 2017; Fathi-Kazerooni et collab., 2019] ou des quelques premiers paquets d'un flux [Lopez-Martin et collab., 2017; Salman et collab., 2018]. Cependant, ces approches nécessitent d'identifier des caractéristiques prédéfinies basées sur les flux avec une intervention humaine, ce qui nécessite donc une base de connaissances qui concerne les applications considérées et un temps d'analyse important. Par conséquent, de nombreuses études se concentrent sur une approche de classification du trafic utilisant les approches d'apprentissage profond. L'objectif de l'algorithme DL est de combiner les modules d'extraction de caractéristiques et de classification en un seul module, de sorte que cette approche ne nécessite pas d'intervention humaine pour les extractions de caractéristiques. Cette méthode convertit la charge utile du paquet en octets pour créer un vecteur représentant une classe d'application. Ensuite, les algorithmes DL analysent ce vecteur pour extraire automatiquement les caractéristiques basées sur les paquets (caractéristiques implicites) et classer ces caractéristiques dans les classes d'application pour le trafic HTTPS [Wang et collab., 2018] et VPN [Wang et collab., 2017; Lotfollahi et collab., 2020].

De nombreuses études existantes se concentrent sur les solutions de classification de trafic pour HTTPS [Wang et collab., 2018] et le trafic VPN [Wang et collab., 2017; Lotfollahi et collab., 2020]. Dans ce travail, nous nous concentrons sur une solution de classification de trafic pour le trafic QUIC. D'après les expériences, l'analyse des données utiles des paquets QUIC illustre que les approches de classification de trafic basées sur le DL ne peuvent pas être efficaces en cas de coexistence de plusieurs applications, y compris les applications de chat et de VoIP. La raison en est qu'il y a moins de différences pour certaines applications telle que la messagerie instantanée (chat) et la VoIP.

4.3.2 Une nouvelle méthode de classification du trafic basée sur le réseau de neurones convolutifs

Afin de répondre au besoin des opérateurs d'identifier le type trafic y compris dans le cadre d'une communication chiffrée et particulièrement pour des communications basées sur le protocole QUIC, nous avons proposé une nouvelle méthode de classification du trafic utilisant des caractéristiques basées sur les flux ainsi que sur des caractéristiques basées sur les paquets. On considère deux types de flux réseau : les flux de souris (un petit flux continu) et les flux éléphants

(un flux continu extrêmement volumineux). Par conséquent, la proposition comprend deux étapes principales de classification. Dans la première étape, les caractéristiques basées sur les flux sont analysées par des algorithmes ML pour identifier les applications de chat et de VoIP (flux de souris). Ensuite, les caractéristiques des flux d'éléphants basées sur les paquets sont extraites et analysées par un réseau de neurones convolutifs (CNN) pour classer les flux d'éléphants en 3 catégories : transfert de fichiers, streaming vidéo ou Google Play Music (flux audio).

Collecte de trafic

Dans la classification du trafic, nous avons besoin de collecter le trafic réseau et de l'analyser pour identifier les classes d'applications. Le trafic réseau peut être collecté directement à partir des commutateurs Openflow dans l'environnement SDN, mais cela n'est pas efficace et s'avère très consommateur de ressources. Par conséquent, il est nécessaire de mettre en place une stratégie qui prend en charge la technique d'échantillonnage pour réduire le volume du trafic. Il existe deux techniques d'échantillonnage : l'échantillonnage de flux et l'échantillonnage de paquets. L'échantillonnage de flux (par exemple, NetFlow, etc.) échantillonne un flux avec une probabilité spécifique, regroupe tous les paquets de ce flux dans des enregistrements et les exporte vers des collecteurs. L'échantillonnage de paquets (par exemple, sFlow, etc.) échantillonne chaque paquet avec une probabilité spécifique et regroupe les paquets échantillonnés dans des enregistrements et les exporte vers des collecteurs. La technique d'échantillonnage de flux peut collecter trop de paquets de certains flux et aucun paquet pour d'autres flux. Par conséquent, elle est inefficace pour la classification du trafic. Nous avons donc décidé d'adopter une stratégie d'échantillonnage de paquets basée sur sFlow [Ujjan et collab., 2020].

sFlow contient deux éléments principaux : les agents sFlow et les collecteurs sFlow. Les agents sFlow sont mis en œuvre au niveau des nœuds de bordure du réseau. Lorsque du trafic réseau traverse les agents sFlow, ces derniers échantillonnent ces paquets, les agrègent et les exportent vers les collecteurs pour la classification du trafic.

Approche proposée pour la classification du trafic

La figure 4.1 décrit l'ensemble de la méthode de classification du trafic proposée. Tout d'abord, le trafic réseau est collecté, puis les premiers paquets de chaque flux sont traités dans le module d'extraction de caractéristiques pour extraire huit caractéristiques basées sur le flux représentant le flux réseau. Plus on utilise de paquets dans un flux, plus on peut obtenir une précision élevée. Cependant, cela entraîne un temps de traitement élevé. Par conséquent, nous considérons les vingt premiers paquets de chaque flux pour extraire les caractéristiques basées sur le flux comme dans l'étude existante [Lopez-Martin et collab., 2017]. Ensuite, ces caractéristiques sont analysées par des algorithmes d'apprentissage automatique qui classent les flux du réseau en flux de chat, de VoIP et flux éléphants. Il existe de nombreux algorithmes de classification, mais l'algorithme de forêt aléatoire s'est avéré être le plus performant [Zhou et collab., 2020]. Après avoir identifié les applications de chat et de VoIP, le reste du trafic réseau (flux éléphants) est traité dans la deuxième étape de classification pour classer les flux éléphants en streaming vidéo, transfert de fichiers et application Google Play Music.

Tout d'abord, le paquet de flux éléphants est traité dans le module de prétraitement pour extraire un vecteur de 1400 valeurs décimales représentant chaque paquet. Ensuite, ce vecteur est analysé dans le module de classification multi-classes pour identifier la classe d'application. La structure de ce module est décrite dans la Fig. 4.2. La taille du vecteur d'entrée étant importante (1400 valeurs), les ANN ainsi que les algorithmes traditionnels (par exemple, SVM, MLP, etc.) ne sont pas efficaces pour le traitement de ces données [Goodfellow et collab., 2016]. En outre, CNN contient des caractéristiques telles que la connectivité de réserve, le partage des paramètres et les représentations équivariantes. Cela offre des possibilités d'apprendre des représentations plus efficaces par rapport aux algorithmes d'apprentissage automatique traditionnels [Goodfellow et collab., 2016]. Par conséquent, l'algorithme CNN est pris en compte dans le module de classification.

La structure du module de classification multi-classes contient cinq couches essentielles, notamment la couche de convolution, la couche de mise en commun des moyennes, la couche de mise à plat, la couche connectée et la couche softmax. La couche de convolution est considérée comme un apprentissage de caractéristiques pour extraire les caractéristiques basées sur les paquets représentant un paquet dans les flux d'éléphants. La couche de mise en commun des moyennes remplace le vecteur d'entrée à un certain endroit par une valeur moyenne des valeurs voisines, tandis que la couche de mise à plat permet de mettre bout à bout toutes entrées sous la forme d'un vecteur unidimensionnel. Dans la couche connectée, chaque neurone a des connexions complètes avec chaque neurone des couches suivantes. La couche softmax est une généralisation d'une fonction logistique qui écrase un vecteur n-dimensionnel en un vecteur n-dimensionnel avec des valeurs d'éléments entre 0 et 1. La combinaison de la couche connectée et de la couche softmax est similaire à un classificateur qui indique une relation entre les vecteurs d'entrée et les classes d'application pour classer les flux éléphants en différents types d'applications.

Dans chaque flux éléphant, nous extrayons des caractéristiques des premiers paquets, classons ces derniers et agrégeons les résultats de la classification afin d'identifier la classe d'application selon la règle de majorité. La règle de majorité est une règle de décision qui sélectionne une alternative avec des votes élevés. Si la plupart des paquets sont classés comme une application donnée, le flux sera attribué à cette dernière. Dans le cas d'une égalité, une probabilité de classification de ces paquets sera utilisée pour identifier la classe d'application. Plus le nombre de paquets classés dans chaque flux est élevé, plus la précision est grande. Néanmoins, cela peut conduire à un temps de traitement élevé. Par conséquent, nous considérons les dix premiers paquets de chaque flux d'éléphants pour identifier la classe d'application. Le détail de la règle de majorité est décrit dans l'algorithme 1.

Algorithm 1 Majority rule

Require: $id, FlowId, label, FlowLabel, service, Count, prob, Prob, a$

```

1: while  $id$  in  $FlowId$  do
2:   for  $j = 1$  to 3 do
3:     for  $i = 1$  to 10 do
4:       if  $label(id)$  is  $service_j$  then
5:          $a_{ij} = 1$ 
6:       else
7:          $a_{ij} = 0$ 
8:       end if
9:     end for
10:  end for
11:  if all  $Count_j$  unique in  $Count$  then
12:     $FlowLabel(id) = \text{argmax}(Count)$ 
13:  else
14:     $FlowLabel(id) = \text{argmax}(Prob)$ 
15:  end if
16: end while

```

$$Count_j = \sum_{i=1}^{10} a_{ij}$$

$$Prob_j = \sum_{i=1}^{10} prob_{ij}$$

$label(id)$ est une classe d'application du flux après la classification et $prob_{ij}$ est une probabilité que le paquet $packet_i$ soit classé comme application app_j dans les flux. Argmax est une fonction qui renvoie un indice de la valeur la plus élevée dans l'entrée.

La proposition contient deux algorithmes principaux, à savoir la forêt aléatoire et le CNN. Par conséquent, la complexité temporelle de la proposition est liée à la complexité temporelle de ces algorithmes. En ce qui concerne l'algorithme de forêt aléatoire, sa complexité temporelle dans la

phase d'apprentissage est de $O(n_0 * \log(n_0) * d * k)$ où n_0 est le nombre d'échantillons d'entrée, d est la dimension des données et k est le nombre d'arbres dans la forêt tandis que la complexité temporelle dans la phase de test est de $O(k * d')$ où d' est la profondeur de l'arbre.

Quant au CNN, la complexité temporelle totale de toutes les couches convolutionnelles [He et Sun, 2015] est estimée par Equ. 4.1 :

$$O\left(\sum_{l=1}^d n_{l-1} \cdot s_l^2 \cdot n_l \cdot m_l^2\right). \quad (4.1)$$

où d est le nombre de couches convolutives, n_{l-1} est le nombre de canaux d'entrée de la l e couche, n_l est le nombre de filtres dans la l e couche, s_l est la taille spatiale du filtre et m_l est la taille spatiale de la carte de caractéristiques de sortie.

Cette complexité temporelle est appliquée à la fois à la phase de formation et à la phase de test, et le temps de formation est environ trois fois supérieur au temps de test.

4.3.3 Validation expérimentale

La qualité de la classification globale peut être évaluée de deux façons : Micro-Moyenne et Macro-Moyenne [Sokolova et Lapalme, 2009]. Dans le cas d'une macro-moyenne, la moyenne d'une métrique est calculée sur toutes les classes qui sont traitées de manière égale. La micro-moyenne est basé sur le cumul des vrais positifs (TP), faux positifs (FP), vrais négatifs (TN) et faux négatifs (FN) de l'ensemble de données.

Les performances de la proposition sont évaluées en utilisant 3 mesures, à savoir la précision, le rappel et le score F1 (mesure F). La précision est un pourcentage des flux pertinents qui sont récupérés, tandis que le rappel est un pourcentage des flux récupérés qui sont pertinents. Le score F1 représente une moyenne harmonique entre la précision et le rappel. Le détail de ces paramètres est décrit dans l'équation : 4.2, 4.3, 4.4.

$$\text{Précision} = \frac{\text{TP}}{\text{TP} + \text{FP}}. \quad (4.2)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}. \quad (4.3)$$

$$\text{F1-score} = \frac{2}{1/\text{Rappel} + 1/\text{Précision}}. \quad (4.4)$$

Tout d'abord, nous utilisons l'algorithme de forêt aléatoire et les caractéristiques basées sur le flux pour identifier les services Google Hangout contenant une discussion et un appel vocal. Les performances des services de chat et d'appel vocal sont supérieures à 96 %. Cependant, le rappel du service de chat est le plus faible avec 96,54 %. Les services de chat et d'appel vocal ont des structures similaires au niveau paquets, de sorte que certains flux du service d'appel vocal sont identifiés dans le service de chat. Cela conduit à un résultat plus faible pour le service de chat. En particulier, les performances du troisième groupe (transfert de fichiers, streaming vidéo et Google Play Music) sont supérieures à 99,5 %. La raison en est qu'il y a de légères différences dans les caractéristiques basées sur le flux du service de chat, des services d'appels vocaux et du troisième groupe. Par conséquent, les moyennes micro et macro de la précision, du rappel et du f1-score atteignent plus de 98 %.

Ensuite, le reste de l'ensemble de données est étudié dans deux scénarios pour détecter le transfert de fichiers, le streaming vidéo ou Google Play Music. Dans le premier scénario, nous étudions l'influence de différentes caractéristiques sur notre méthode proposée. Nous mettons en œuvre cinq ensembles de données de caractéristiques possibles et apprécions les caractéristiques qui indiquent une haute performance. La Fig. 4.3 indique cinq jeux de données comprenant les 300, 600, 900, 1200 et 1400 premières caractéristiques dans les données utiles cryptées des paquets QUIC. Il y a une tendance à la hausse dans la précision, le rappel et le score f1 des micro et macro-moyennes dans les cinq ensembles de données. Les performances avec les 300 et 900 premières

Class	Precision	Recall	F1-score	Nombre de flux
Chat	0,9871	0,9654	0,9762	2,227
Appel vocal	0,9871	0,9880	0,9875	2,086
Troisième groupe	0,9957	0,9997	0,9977	11,351
Micro-moyennage	0,9934	0,9934	0,9934	15,664
Macro-étalonnage	0,9899	0,9843	0,9871	15,664

TABLEAU 4.1 – Précision, rappel et score F1 pour la première étape de la classification à l'aide de l'algorithme de forêt aléatoire

caractéristiques ne sont pas bonnes (de l'ordre de 70% dans la moyenne micro et macro du f1-score). La précision du service de streaming vidéo avec les 300 premières caractéristiques est de 64,34 %. De même, certains flux du service de streaming vidéo sont classés dans les flux du service de transfert de fichiers, ce qui réduit le rappel du service de transfert de fichiers (seulement 26,52 %). C'est similaire à l'ensemble de données des 900 premières caractéristiques. En outre, la performance avec l'ensemble de données de 1400 caractéristiques est la plus élevée avec plus de 97 pour cent dans cinq types de l'ensemble de données. Le pouvoir discriminant de 1 400 caractéristiques est plus important que celui des autres caractéristiques, de sorte que la précision, le rappel et le score f1 des micro et macro-moyennes sont les plus élevés, avec environ 99 %.

Dans le second scénario, nous étudions l'influence de certaines fonctions de perte [CholletKeras, 2016] sur le réseau neuronal convolutif (Fig. 4.4). La performance de l'ensemble de données avec 1400 caractéristiques est la plus élevée parmi les cinq types d'ensembles de données, nous utilisons donc cet ensemble de données dans le deuxième scénario. Nous mettons en œuvre trois types de fonctions de perte : categorical_hinge (Hinge), mean_squared_error (MSE) et sparse_categorical_crossentropy (SCCE). Dans ce scénario, nous évaluons les performances de trois types de fonctions de perte afin de sélectionner la fonction de perte appropriée pour notre méthode. Hinge Loss est la fonction de perte qui est notamment utilisée dans les machines à vecteurs de support (SVM) pour la classification à "marge maximale". Dans la Fig. 4.4, la performance de la Hinge Loss n'est pas la meilleure métrique parce que les micro et macro-moyennes de précision, le rappel et le score f1 sont supérieurs à 20 %. En particulier, les flux de transfert de fichiers ne peuvent pas être détectés. L'erreur quadratique moyenne (MSE) est la fonction de perte qui mesure la moyenne des carrés des erreurs. La performance de la MSE est supérieure à celle de la Hinge Loss. De plus, les flux des services de streaming vidéo sont inférieurs à ceux des autres services. La raison en est que les flux de streaming vidéo sont traités comme les flux de Google Play Music, de sorte que le rappel de Google Play Music n'est que de 94,73

Grâce à deux scénarios, nous avons constaté que l'ensemble de données avec 1400 caractéristiques et la fonction de perte SCCE sont appropriés pour notre méthode proposée. La Fig. 4.5 indique les résultats globaux de nos méthodes proposées avec l'ensemble de données de 1400 caractéristiques et la fonction de perte SCCE sur cinq types de services basés sur QUIC. Les moyennes micro et macro de la précision, du rappel et du score f1 sont d'environ 99 %. Ces résultats sont calculés en se basant sur la boîte à outils scikit-learn [Pedregosa et collab., 2011].

En conclusion, cette contribution avait pour objectif de proposer une méthode d'identification du trafic. L'objectif étant de proposer la meilleure expérience utilisateur possible en fonction du service utilisé. Par conséquent, la contribution suivante va s'intéresser à la mesure qualitative de cette expérience.

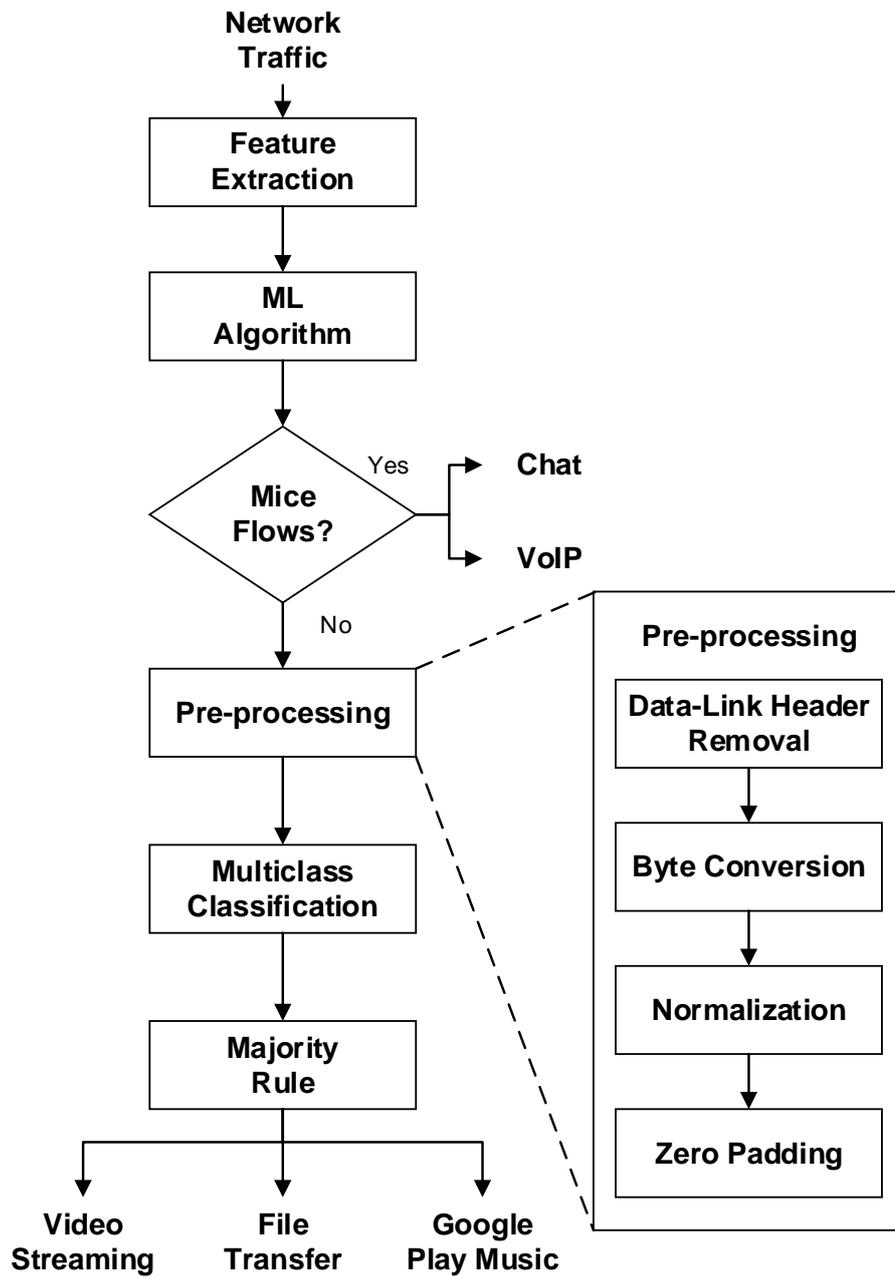


FIGURE 4.1 – Structure générale

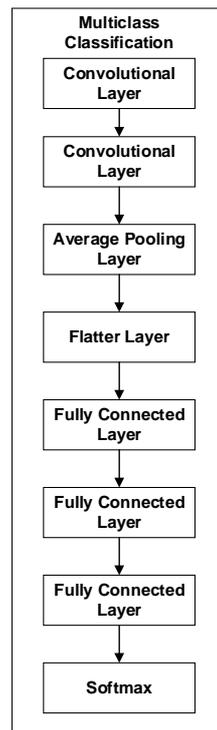


FIGURE 4.2 – Multi-Classifier

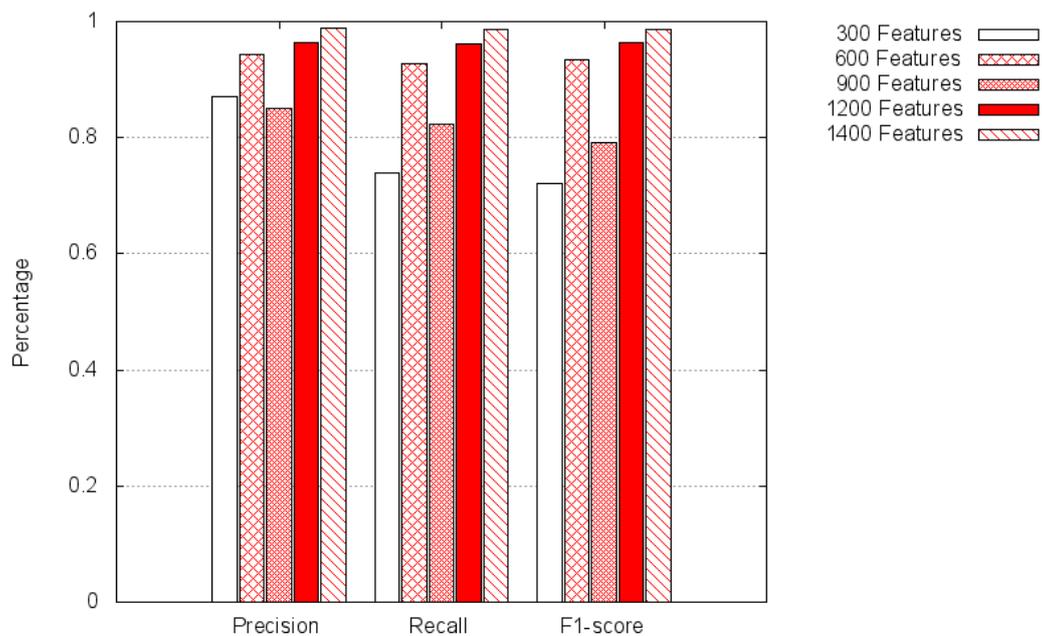


FIGURE 4.3 – Précision macro-moyenne, rappel macro-moyenne et f1-score macro-moyenne dans différents ensembles de données

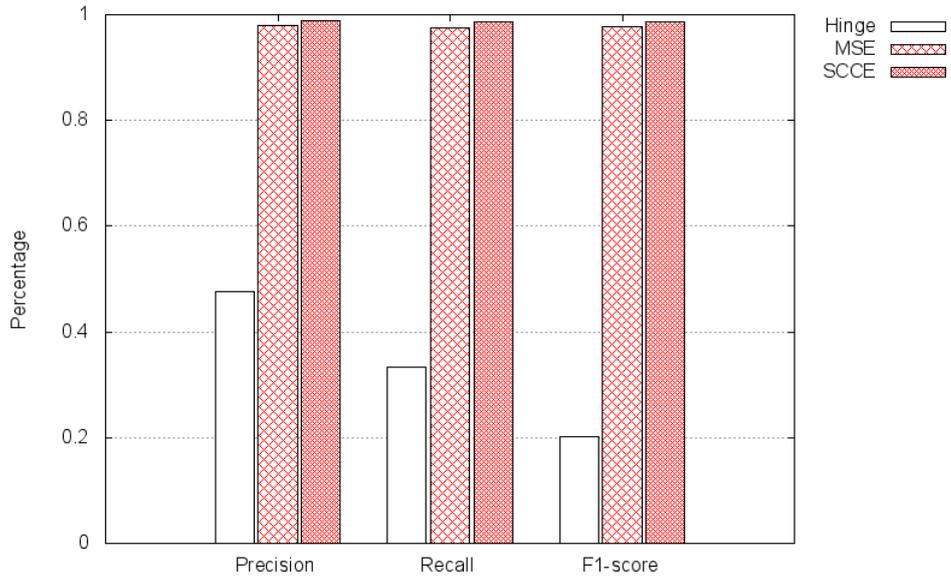


FIGURE 4.4 – Macro-moyenne de la précision, macro-moyenne du rappel et macro-moyenne du f1-score pour différentes fonctions de perte

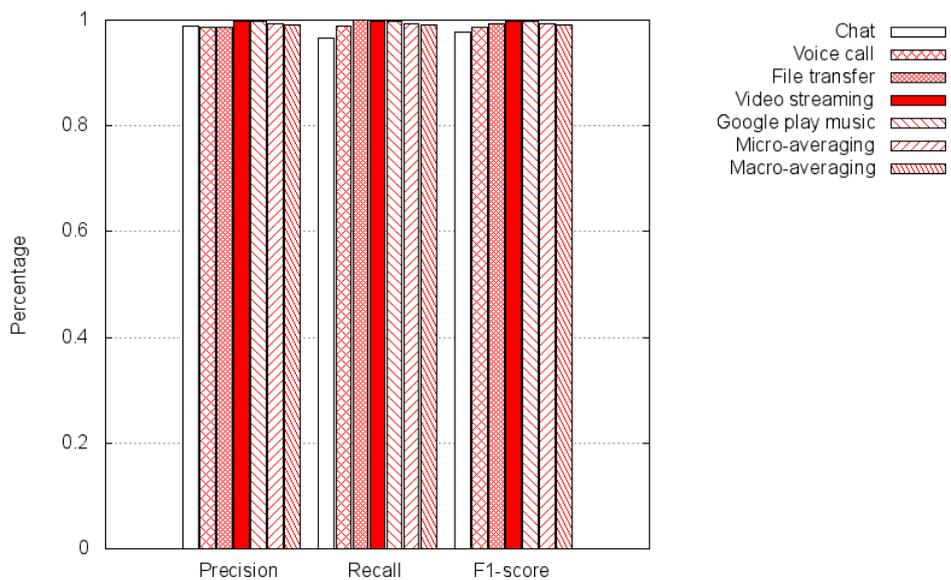


FIGURE 4.5 – Précision, Recall et F1-score de notre méthode proposée sur cinq types de services basés sur QUIC

4.4 Contribution 2 : Estimation de la QoE

Dans un second temps, nous avons travaillé sur la notion de qualité telle qu'elle était perçue par l'utilisateur, autrement dit, la qualité d'expérience (QoE). La notion de QoE est apparue au cours de la dernière décennie avec l'intention d'impliquer la perception subjective des utilisateurs finaux dans l'évaluation de la qualité des services de télécommunication [Aroussi et Mel-louk, 2014]. Aujourd'hui, il est particulièrement difficile d'établir une définition claire et précise du concept de qualité d'expérience en raison des nombreux facteurs/domaines qui influencent l'expérience de l'utilisateur. Néanmoins, nous présentons quelques définitions qui, selon nous, reflètent le mieux la notion de qualité d'expérience. Selon la normalisation des télécommunications¹ (ITU-T), la qualité d'expérience est définie comme "*l'acceptabilité globale d'une application ou d'un service, telle qu'elle est perçue subjectivement par l'utilisateur final*". [ITU, pages 1-8. January 2006]. Du point de vue de l'institut européen des normes de télécommunications (ETSI), la qualité d'expérience de l'utilisateur illustre "*une mesure de la performance de l'utilisateur basée sur des mesures psychologiques objectives et subjectives de l'utilisation d'un service ou d'un produit des technologies de l'information et de la communication (TIC)*". [ETS, November 2009]. Une autre définition est donnée par le réseau européen sur la qualité de l'expérience dans les systèmes et services multimédias (QUALINET) où la QoE représente le "*degré de plaisir ou d'ennui que vous éprouvez à l'égard d'un produit, d'une application ou d'un service*". [P. Le Callet et Perkis, March 2013].

Il existe différentes mesures qui ont un impact direct/indirect sur la QoE ressentie par l'utilisateur final. Ces mesures sont appelées *facteurs d'influence de la qualité d'expérience (QoEIF)* [Ickin et collab., 2012]. Ces facteurs sont étroitement liés à la perception humaine et peuvent potentiellement servir d'indicateurs de qualité pour les fournisseurs de services. Cependant, certains problèmes apparaissent de manière relativement récurrente dans l'évaluation de la qualité d'expérience. Le premier problème est la différence de retour d'informations entre les utilisateurs (divergences d'opinion pour le contexte), et le second consiste en la relation de non-linéarité entre la QoE et les facteurs d'influence. Bien que différentes méthodes (objectives, subjectives, hybrides) puissent être utilisées pour l'évaluation de la qualité d'expérience des utilisateurs [Liotou et collab., 2016; Reichl et collab., 2010; Mushtaq et collab., 2016; Alreshoodi et Woods, June 2014], une méthode d'évaluation de la qualité d'expérience plus efficace est toujours nécessaire.

Notre contribution a consisté à proposer une nouvelle méthode d'évaluation de la qualité d'expérience appelée modèle *Belief Functions Combination (BFC)* pour mesurer la qualité d'expérience réelle telle que perçue par les utilisateurs. Nous nous sommes concentrés sur le trafic vidéo dans un premier temps. Cette méthode utilise la théorie des fonctions de croyance et les règles de fusion des combinaisons pour calculer l'expérience de l'utilisateur en termes d'échelle d'évaluation du score d'opinion moyen (MOS) [ITU, 2016]. Elle fusionne les différentes valeurs d'estimation données par diverses sources d'information, que nous appelons "oracles" et qui sont basées sur plusieurs méthodes d'apprentissage automatiques (ML).

4.4.1 Estimation de la QoE via des techniques de ML

De nombreuses méthodes d'apprentissage automatique ont été appliquées avec succès pour prédire la qualité d'expérience des utilisateurs dans des environnements contrôlés et non contrôlés. Dans nos travaux, nous avons considérés 6 modèles de ML pour l'estimation de la QoE.

- Les machines à vecteurs de support (SVM) : il s'agit d'une méthode de classification qui analyse un échantillon de données et tente d'identifier des modèles afin de classer ces données. L'idée principale est de trouver l'hyperplan de séparation optimale entre les classes en maximisant la marge entre les points les plus proches [H. Mark and F. Eibe and H. Geoffrey and P. Bernhard and R. Peter and W. Ian, June 2009].

1. <https://www.itu.int/en/ITU-T/Pages/default.aspx>

- Les réseaux neuronaux aléatoires (RNN) sont une variante des réseaux neuronaux. Il s'agit d'une représentation mathématique d'un réseau interconnecté de neurones, qui échangent périodiquement des signaux. Dans les RNN, chaque neurone a un potentiel à actualiser. Si le potentiel est positif, alors le neurone peut transmettre des signaux, sinon il ne fait que mettre à jour son potentiel. En fait, les signaux échangés ont deux états : excité ou inhibé. Lorsqu'un signal est excité, il augmente le potentiel du neurone qui l'a reçu. Dans le cas d'un signal inhibiteur, il réduit le potentiel du neurone.
- Les méthodes basées sur le bagging ont été introduites par Breiman. Il s'agit d'entraîner différents classificateurs avec des répliques bootstrappées de l'ensemble de données d'entraînement original. Cela signifie qu'un nouvel ensemble de données est formé pour entraîner chaque classificateur en tirant au hasard (avec remplacement) des instances de l'ensemble original des données (généralement, en conservant la taille de l'ensemble original des données). Par conséquent, une diversité est obtenue avec la procédure de ré-échantillonnage par l'utilisation de différents sous-ensembles de données [M. Galar and A. Fernandez and E. Barrenechea and H. Bustince and F. Herrera, 2012]. Dans ce travail, nous avons considéré la méthode Random Forest ("RF") [Aydin et collab., 2016] et la méthode d'arbre de décision avec bagging ("DTbagging") [J. Mendes-Moreira et Sousa, 2012].
- La méthode de boosting est un moyen d'améliorer les performances de tout algorithme d'apprentissage [Freund et Schapire, pages 148–156. July 1996]. Le processus de boosting fonctionne en deux étapes. Dans la première étape, il utilise des sous-ensembles de données originaux pour produire une série de modèles modérément efficaces. Dans la deuxième étape, il amplifie ses performances en les combinant à l'aide d'une fonction de coût particulière (vote majoritaire). Contrairement au Bagging, dans la stimulation classique, la création d'un sous-ensemble n'est pas aléatoire, mais dépend des performances des modèles précédents : chaque nouveau sous-ensemble contient les éléments qui ont été (probablement) mal classés par les modèles. Selon [P. K. Mallapragada and R. Jin and A. K. Jain and Y. Liu, 2009], les méthodes basées sur le boosting peuvent réduire l'erreur de tout algorithme d'apprentissage, puisqu'il suffit d'être meilleur qu'une supposition aléatoire. Dans ce travail, nous avons considéré deux méthodes de boosting : le boosting basé sur le gradient ("Boost-Grad") [Chen et collab., May 2013] et le boosting basé sur l'arbre de décision ("BoostDT") [Tantithamthavorn et collab., 2016].

Comme énoncé précédemment, on s'est principalement focalisé sur les services vidéos. Dans le cas de la diffusion d'un flux vidéo, il existe un grand nombre d'indicateurs et de paramètres qui peuvent influencer la qualité de la vidéo (par exemple, le taux de bits de la vidéo, le taux d'images de la vidéo, etc.), mais leur effet combiné n'est pas bien identifié. C'est ce que nous appelons : "QoE *IFs interaction problem*". Pour résoudre ce problème et améliorer les performances de l'évaluation de la qualité d'expérience, il est nécessaire de réaliser une étape de sélection des caractéristiques afin de choisir les paramètres appropriés à utiliser par les méthodes ML. Cette procédure de sélection peut être réalisée à l'aide d'une analyse par composantes principales (ACP) [Khan et collab., 2009; Youssef et collab., December 2016] et via une matrice de corrélation de Pearson [L. Amour, S. Souihi, M. S. Mushtaq, S. Hoceini and A. Mellouk, 2016]. Dans un premier temps, nous utilisons 11\$ paramètres vidéo collectés à partir de l'ensemble de données subjectives collectées pour découvrir les facteurs à forte influence qui contribuent à la qualité d'expérience de l'utilisateur. La matrice de corrélation entre les différents paramètres est présentée dans figure 4.6).

Sur la base de la matrice de corrélation, nous mettons en œuvre une méthode heuristique [L. Amour, S. Souihi, M. S. Mushtaq, S. Hoceini and A. Mellouk, 2016] pour éviter les problèmes d'interaction et pour sélectionner les meilleurs paramètres pour l'estimation de la QoE. Les caractéristiques sélectionnées sont : la résolution vidéo (*Resol*), la fréquence d'images vidéo (FR), la perte d'images vidéo (FL), le débit binaire vidéo (BR) et le débit binaire audio (AR). Sur la base de ces paramètres nous analysons la performance de prédiction (entre le MOS prédit et le MOS réel) à l'aide de deux métriques, comme le suggère [VQE, 2005. Access : www.vqeq.org; A. Khan, L. Sun and E. Ifeachor, 2010] : (i) MAE (équation 4.5) et (ii) la corrélation "r" (équation 4.6).

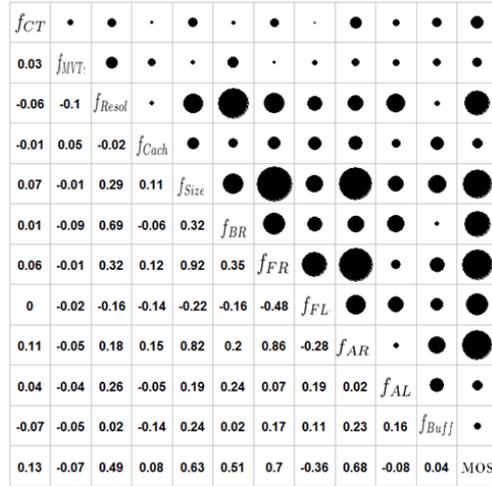


FIGURE 4.6 – Corrélation de la QoE avec les paramètres vidéo

$$MAE = \frac{1}{nb} \sum_1^{nb} |MOS_{real} - MOS_{Model}| \quad (4.5)$$

où : MOS_{Model}^n est la valeur prédite à l'aide de d'une méthode ML donnée, MOS_{real} est la valeur réelle et nb est le nombre total des échantillons considérés.

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}} \quad (4.6)$$

où :

- " y_i " présente le score d'évaluation MOS réel; " x " présente les valeurs prédites du score MOS et " n " présente le nombre d'échantillons.

Nselon l'équation

ID	Nom du modèle	MAE	r
SVM	Support Vector Machine	0.528	0.842
RNN	Réseau de neurones aléatoires	0.623	0.831
RF	Random Forest	0,270	0,894
BaggingDT	Bagging avec DT	0,321	0,872
BoostDT	Bossting avec DT	0.274	0.887
BoostGrad	Boosting avec Gradient	0.244	0.904

TABEAU 4.2 – Tableau de comparaison des méthodes testées

D'après notre évaluation ML, on observe que "*BoostGrad*" est le meilleur modèle de prédiction la QoE de l'utilisateur dans notre cas. La figure (4.7), représente la distribution de la meilleure méthode prédictive pour chaque échantillon [Amour et collab., October 2017].

Nous observons que dans 74% des cas environ, la méthode "*BoostGrad*" donne la valeur MOS prédite la plus proche de celle fournie par l'utilisateur dans l'environnement. Cependant, d'autres méthodes de ML ("*SVM*", "*RNN*", "*RF*" et "*BaggingDT*") fournissent de meilleures performances de prédiction dans 26% des cas. Nous avons conclu qu'il peut être intéressant de combiner ces méthodes afin d'obtenir la meilleure valeur prédite du MOS en fonction du contexte. En d'autres

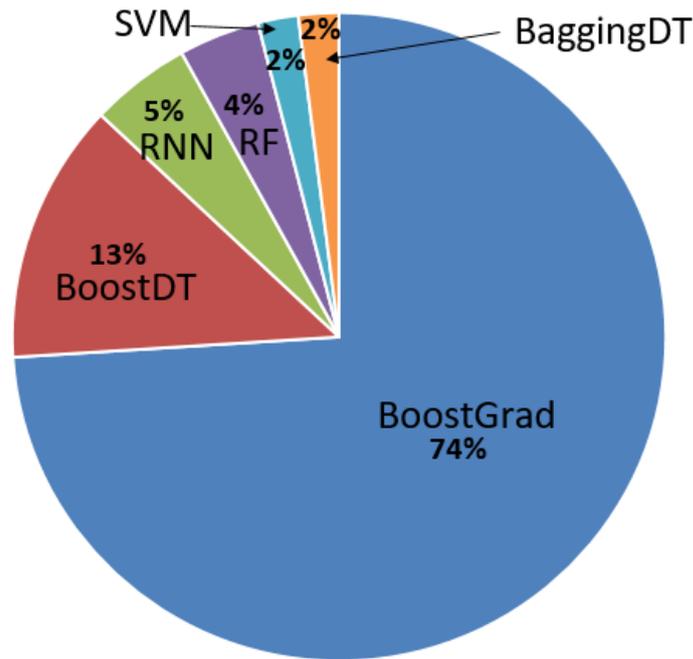


FIGURE 4.7 – Distribution de la meilleure méthode de prédiction pour chaque échantillon

termes, il s'agit de trouver la meilleure valeur MOS prédite de la qualité d'expérience de l'utilisateur en combinant toutes les valeurs MOS prédites calculées par les méthodes ML ou en choisissant la meilleure.

4.4.2 Nouvelle approche d'estimation de la QoE

Partant du constat qu'il est possible de combiner plusieurs techniques d'estimation de la QoE, nous avons proposé une méthode appelée *Belief Functions Combined (BFC)*, qui fusionne les valeurs calculées fournies par diverses méthodes ML en utilisant la théorie des fonctions de croyance et les règles de fusion des combinaisons (figure (4.10)).

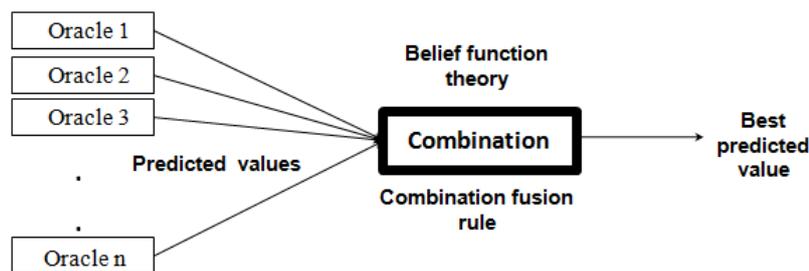


FIGURE 4.8 – Présentation générale du modèle de prédiction Belief Functions Combination (BFC)

La théorie des fonctions de croyance a été introduite par [G. Shafer \[Princeton University Press, pages 1-314. April 1976\]](#). Cette théorie a été développée par Philippe Smets pour aboutir au *Modèle de croyance transférable (TBM)* [Smets, 1998]. Ce modèle est un cadre formel pour la représentation et la combinaison des connaissances. Il est basé sur la définition de fonctions de croyance fournies par des sources d'information, appelées "oracles", qui peuvent être élémentaires, redondantes ou potentiellement non-indépendantes. Elle propose également un ensemble d'opérateurs permettant de combiner ces fonctions. En pratique, le TBM peut donc naturellement être utilisé dans le contexte de la fusion d'informations pour améliorer l'analyse et l'interprétation de données provenant d'oracles multiples.

Contrairement à d'autres modèles de croyances (le modèle de croyances de [A. P. Dempster \[325-339. February 1967\]](#) et le modèle basé sur les probabilités [[P. Walley, pages 1-700. December 1991](#)]), le modèle TBM est caractérisé par son caractère fondamentalement non probabiliste, où une justification axiomatique cohérente est fournie pour les principaux concepts de la théorie des fonctions de croyance. Cette justification a clarifié le lien entre la représentation des croyances et la prise de décision [[Smets, 1998](#)]. Les mécanismes de raisonnement du TBM sont ainsi regroupés en deux niveaux comme l'illustre la figure (4.9).

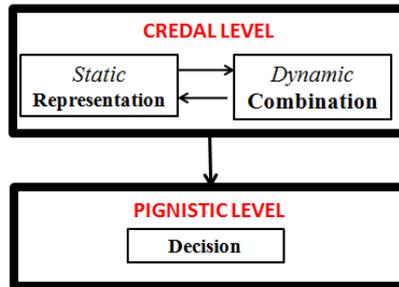


FIGURE 4.9 – Description du modèle de croyances transférable (TBM)

Le premier niveau appelé "*crédal*" concerne la représentation des connaissances (partie statique) ainsi que les combinaisons et le raisonnement sur ces connaissances (partie dynamique). Le second niveau dit "*pignistique*" concerne la prise de décision, en tenant compte du risque et/ou du gain associé à cette décision. En pratique, le niveau crédal représente la modélisation des croyances sous forme de fonctions de croyance, et le niveau pignistique consiste à transformer ces croyances en probabilités pour prendre des décisions. Ces deux niveaux sont présentés ci-dessous.

L'architecture globale du modèle proposé, appelé *Belief Functions Combined (BFC)*, se compose de trois modules individuels, l'un pour la sélection des paramètres, le second pour l'évaluation des oracles, et le dernier pour la combinaison des résultats en utilisant la théorie des fonctions de croyance. Ce module dernier module est basé sur le modèle TBM qui a été introduit par [Smets \[1998\]](#). Ce module est composé de deux niveaux (Figure 4.9).

Niveau crédal

Dans ce niveau, les croyances sont quantifiées par des fonctions de croyance, appelées "masse", où ces croyances sont entretenues et mises à jour en utilisant deux étapes.

Étape 01 : Description du cadre de discernement (Θ) :

Il représente les décisions possibles que l'oracle peut prendre. Ces valeurs représentent le score MOS (niveau de qualité d'expérience) en termes de valeurs entières comme : $\Theta = \text{MOS} = \{1, 2, 3, 4, 5\}$.

Étape 02 : Définition de la fonction de masse :

Elle présente une distribution de probabilité généralisée sur les éléments et le sous-ensemble de Θ ($2^\Theta \rightarrow [0, 1]$) où :

$$\sum_{A \subseteq \Theta} m(A) = 1 \quad (4.7)$$

La fonction de masse est définie pour chaque oracle du système et joue un rôle clé dans la construction du modèle TBM. Selon [S. Le Hégarat-Masclé, I. Bloch and D. Vidal-Madjar \[1997\]](#), il n'existe pas de solution simple pour définir la fonction de masse. Dans [[A. Martin, A. L. Jousselme and C. Osswald, 2008](#)], les auteurs expliquent qu'il n'existe pas de méthode générique pour définir la fonction de masse à moins qu'un expert ne fournisse directement son opinion sous forme de croyance. De plus, la représentation des valeurs de croyance dépend fortement de la modélisation de l'application, et donc la fonction de masse est spécifique à cette application. Cependant, deux

grandes familles émergent pour définir la fonction de masse malgré la spécificité de l'application. La première concerne la famille des mesures de dissimilarité, où l'on trouve deux approches (l'approche globale introduite par Shafer [G. Shafer, Princeton University Press, pages 1-314. April 1976] et l'approche séparable proposée par Smets [Smets, August 1993; A. Appriou, November 1999]). La deuxième famille utilisée pour définir la fonction de masse concerne les méthodes basées sur la distance. Ce type de méthodes est initié par T. Denoeux [1995] et se concentre sur les tendances centrales, puisque la valeur inconnue (fonction de masse) est calculée en fonction du plus proche du centroïde. En outre, une autre approche heuristique existe dans la littérature. Dans cette approche, le calcul de la fonction de masse est basé sur les valeurs de croyance des singletons et sur la distribution de ces valeurs sur les sous-ensembles qui composent le cadre de discernement.

Niveau pignistique

Dans ce niveau, une transformation pignistique des fonctions de croyance en fonctions de probabilité est réalisée [Smets, 1998]. Elle se compose de deux étapes.

Règle de combinaison : Dans cette étape, l'objectif consiste à synthétiser plusieurs fonctions de masse (m_i) définies pour chaque oracle (ML_i), sous la forme d'une fonction de masse unique (m_{rule}) telle que présentée dans l'équation (9).

$$m_{rule} = Combinaison(m_i) \quad (4.8)$$

Où : - i : Indice entre 1 et le nombre d'oracles utilisés.

- **Combinaison** : L'une des règles de fusion. ([E. Lefevre, June 2002]).

Pour atteindre cet objectif, une grande variété de règles de fusion a existé dans la littérature comme : *Florea rule* [Florea et collab., April 2006], *Règle de Dempster-Shafer* [E. Lefevre, June 2002] et *Règle de Murphy* [C.K. Murphy, July 2000].

Prise de décision :

Dans cette étape, une décision doit être déduite de la connaissance résultante (m_{rule}) pour obtenir la meilleure valeur. Habituellement, le maximum de la crédibilité (Bel), de la plausibilité (pl) ou de la probabilité pignistique ($betP$) est pris sur l'espace des décisions admissibles.

- La fonction de crédibilité est donnée pour tout $X \in 2^\Theta$ (m_{rule}) par :

$$Bel(X) = \sum_{Y \in 2^X, Y \neq \emptyset} m(Y) \quad (4.9)$$

- La fonction de plausibilité est donnée pour tout $X \in 2^\Theta$ par :

$$pl(X) = \sum_{Y \in 2^X, Y \cap X \neq \emptyset} m(Y) = Bel(\Theta) - Bel(X^c) \quad (4.10)$$

où : X^c est le complémentaire de l'ensemble X .

- La probabilité Pignistique est donnée pour tout $X \in 2^\Theta$, avec $X \neq \emptyset$ par :

$$betP(X) = \sum_{Y \in 2^X, Y \neq \emptyset} \frac{|Y \cap X|}{|Y|} * \frac{m(Y)}{1 - m(\emptyset)} \quad (4.11)$$

En résumé, la figure 4.10 détaille et résume le processus de mise en œuvre du modèle BFC.

À partir de la figure (4.10), nous observons que la meilleure valeur de prédiction est obtenue en utilisant l'un des critères de décision (par exemple : Bel , pl ou $betP$) comme présenté dans l'équation (4.12).

$$\begin{aligned} MOS_{best}^{t_k} &= GetDecision(m_{rule}) \\ &= GetDecision(Combinaison(m_i)) \end{aligned} \quad (4.12)$$

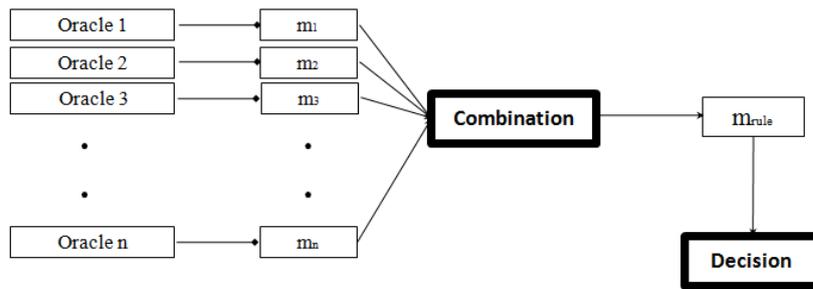


FIGURE 4.10 – Processus d'estimation de la qualité d'expérience

Où :

- **GetDecision** : Critère de prise de décision (par exemple : *Bel, pl* ou *betP*).
- m_i : Fonction de masse construite pour la valeur prédite par le $i^{ième}$ oracle qui représente une méthode ML dans notre modèle de prédiction BFC proposé tel que présenté dans la figure 4.10.

Nous définissons un oracle comme un processus qui prédit la qualité d'expérience en fonction de certains IF de qualité d'expérience. Nous utilisons les méthodes ML comme oracles. En effet, ces méthodes sont souvent utilisées pour évaluer le niveau de QoE (MOS) et ont prouvé leur efficacité dans la pratique [Alreshoodi et Woods, June 2014; Aroussi et Mellouk, 2014].

4.4.3 Évaluation du modèle de prédiction BFC

La figure 4.11, représente les résultats de prédiction en termes de taux d'erreur absolu moyen (MAE) pour la meilleure méthode ML (BoostGrad) et pour chaque règle utilisée dans le modèle BFC. L'axe des abscisses représente le pourcentage de données utilisées dans l'étape d'apprentissage, tandis que l'axe des ordonnées représente le taux d'erreur absolue moyen pour chaque méthode/règle. Cette figure montre qu'à l'exception de la règle conjonctive de Smets ("Smets-BFC") et de la règle de Dempster-Shafer ("DS-BFC"), l'utilisation des règles BF donne de bons résultats pour estimer la qualité en termes de MOS. On observe également qu'en utilisant 10% du jeu de données, la meilleure méthode d'estimation est le modèle BFC utilisant la règle de Murphy pondérée ("Murphy-BFC") avec un MAE de 0,405. On note également que les performances d'estimation s'améliorent en augmentant la taille de l'ensemble de données d'entraînement. Lorsqu'un jeu de données de 80 % est utilisé pour l'apprentissage, nous montrons que la méthode "Murphy-BFC" est la meilleure méthode d'estimation avec une erreur MAE de 0,233.

La figure (4.12) montre les performances de toutes les méthodes d'estimation testées sur la base de l'ensemble du jeu de données de test. L'axe des x représente la meilleure méthode d'estimation ML ("BoostGrad") et cinq règles de combinaison, tandis que l'axe des y représente le taux de MAE pour chaque méthode/règle. Nous observons que les règles de combinaison de base ne sont pas adaptées à l'évaluation de la qualité d'expérience dans notre cas. On observe également que pour la meilleure règle de combinaison, qui est la règle pondérée de Murphy ("Murphy-BFC"), l'erreur absolue moyenne est inférieure à 0,2 , ce qui entraîne une amélioration de 20,5 % par rapport à la méthode "BoostGrad" (meilleure méthode ML).

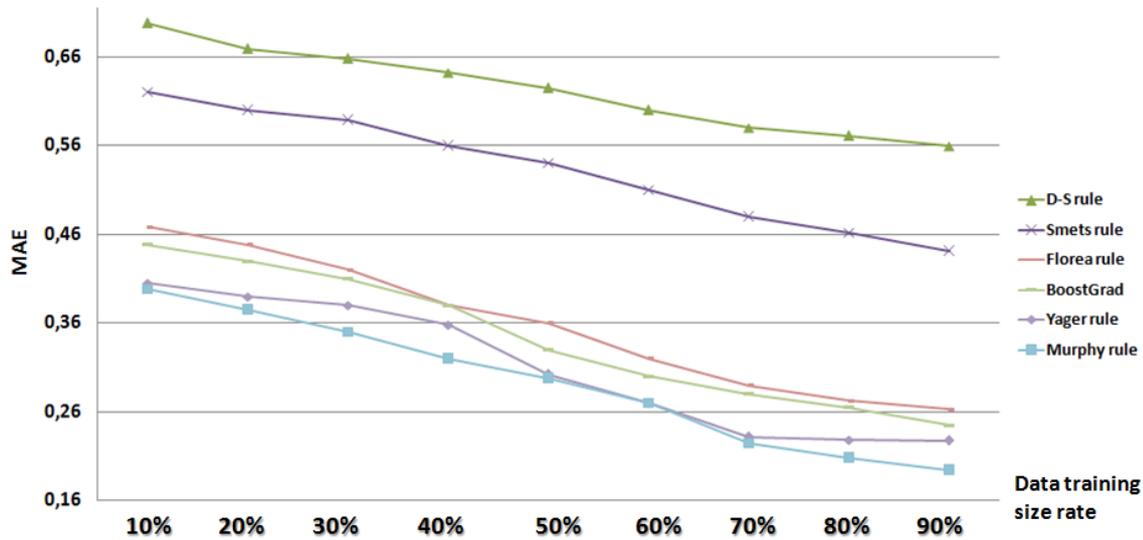


FIGURE 4.11 – Performance de prédiction en termes de taux de MAE lorsque la taille des données d'apprentissage varie.

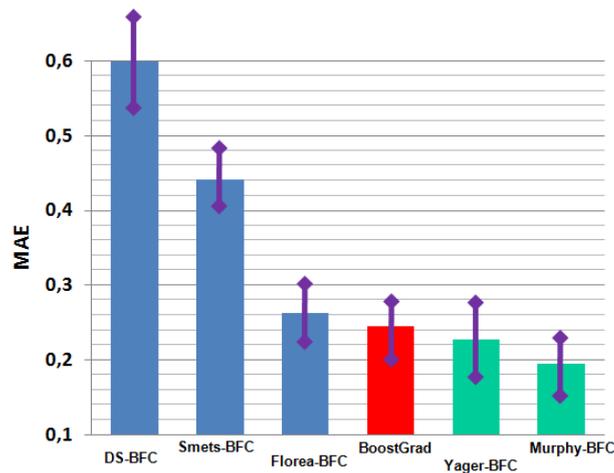


FIGURE 4.12 – Performance de prédiction en termes de taux de MAE

4.5 Conclusion et travail réalisé

Dans la première contribution nous nous sommes concentrés sur l'identification du type de flux pour identifier certains services basés sur QUIC notamment dans un contexte de trafic chiffré. La méthode proposée comporte deux étapes principales de classification : une première étape qui utilise les techniques d'apprentissage automatique classique afin de différencier les flux de souris et les flux éléphants ; la deuxième étape utilise un réseau neuronal convolutif afin de différencier les flux éléphants. Les expériences montrent que la méthode proposée peut détecter cinq types de services basés sur QUIC avec une grande précision (environ 99 %).

Malgré ses performances élevées, la méthode proposée présente quelques inconvénients. L'utilisation des caractéristiques basées sur les flux entraîne une augmentation du temps d'exécution du traitement et de la classification. De plus, la réduction des caractéristiques nécessite un temps de traitement élevé lorsqu'il y a 1400 caractéristiques basées sur les paquets. Ces 2 aspects mériteraient qu'on y consacre davantage d'efforts.

Dans une seconde phase, nous nous sommes intéressés à l'estimation de la QoE. Nous avons étudié l'existant et proposé un modèle de prédiction de la qualité d'expérience, appelé BFC, qui utilise un ensemble de méthodes d'apprentissage automatique (ML) et les fusionne. À cette fin, nous avons utilisé *la théorie des fonctions de croyance* et *les règles de combinaison* afin d'obtenir

un bien meilleur résultat d'estimation. Cette méthode est intéressante car elle est robuste face à un changement de contexte brusque mais néanmoins demeure tributaire des modèles considérés comme Oracle.

4.6 Synthèses

4.6.1 Difficultés rencontrées

L'une des plus importantes difficultés rencontrées dans le cadre de ce travail est probablement le manque de données de qualité. En effet, dans la communauté, il est très rare de trouver en libre accès une base de données de perception de l'utilisateur contenant un grand nombre d'échantillons. Fort heureusement dans le cadre du projet Poqemon nous avons eu l'opportunité de construire une telle base. Néanmoins il est à noter que cette tâche fut très chronophage (environ 2 ans) et a nécessité beaucoup de discussions, d'interrogations, d'échanges et de concertation.

4.6.2 Forces et faiblesses

Selon moi, l'intérêt principal de notre modèle réside dans sa simplicité. En effet, l'approche modulaire que nous avons mis en place permet non seulement d'intégrer de nouveaux modèles développés par la communauté mais aussi de pouvoir s'adapter à d'autres usages.

Néanmoins malgré le fait que ce modèle soit adaptable à d'autres usages il n'en demeure pas moins nécessaire de le nourrir avec des bases de données. Force est de constater qu'à ce jour le partage de ce type de données n'est toujours pas systématique.

4.6.3 Possible extension

Les travaux que nous avons menés sur la QoE peuvent être étendus de deux manières : La première piste consiste à concevoir un module d'estimation de la QoE qui soit agnostique et qui pourrait fonctionner sur tout type de service.

La seconde piste consisterait à prendre l'historique en compte dans l'estimation de la qualité perçue à un instant « t ».

En ce qui concerne la classification du trafic, des mécanismes plus efficaces doivent être élaborés.

4.6.4 Impact potentiel des travaux

On a pris le parti de partager, dès le début de ces travaux, les données que nous avons collectées avec la communauté; l'objectif était de favoriser et de faciliter l'émergence de nouvelles approches plus efficaces à la fois pour l'estimation de la qualité perçue par l'utilisateur et la classification du trafic chiffré. D'un point de vue industriel le fait de développer une approche modulaire permettrait aux ingénieurs de dimensionner le modèle en fonction de la quantité et de la qualité des données dont ils disposent.

4.6.5 Retour d'expérience

Ayant été confronté au début de nos travaux au manque de données nous avons adopté une approche itérative et incrémentale afin de concevoir nos modèles. Les premières contributions ont été de collecter des données de qualité et en quantité qui expriment la perception de l'utilisateur. Sur la base de ces données nous avons entraîné et comparé plusieurs approches d'apprentissage automatique afin de produire un module automatisé d'estimation de la QoE. Ayant constaté qu'en fonction du contexte, certaines approches pouvaient de meilleurs résultats que d'autres, il nous

a paru intéressant de ne pas nous contenter d'un seul estimateur. Nous avons donc étudiés la possibilité d'appliquer la théorie des fonctions de croyance afin de fusionner les résultats obtenus par plusieurs d'entre eux. Il est évident que si nous n'avions pas appliqué une approche incrémentale il aurait été difficile d'arriver aux mêmes conclusions. La même démarche a été suivie pour la problématique de classification du trafic. Nous devons aussi noter l'intérêt et l'importance de mener ce type de travaux de recherche, à dimension plurielle, dans le cadre de projets collaboratifs et de partager les données et les résultats en accès libre avec la communauté scientifique.

Bibliographie

2005. Access : www.vqeq.org, «Video Quality Experts Group, Multimedia Group Test Plan, Draft version 1.8», . 62
- 2016, «International Telecommunication Union - Telecommunication (ITU-T)-T Recommendation P.800.1 : Methods for Objective and Subjective Assessment of Speech and Video Quality : Mean Opinion Score (MOS) Terminology», , p. 1–18. 61
- November 2009, «Human Factors (HF) : Quality of Experience (QoE) Requirements for Real-time Communication Services», dans *ETSI TR 102 643*, p. 1–37. 61
- pages 1-8. January 2006, «ITU-T Recommendation P.10/G.100 : Definition of Quality of Experience (QoE)», . 61
- A. Appriou. November 1999, «Multisensor Signal Processing in the Framework of the Theory of Evidence», *In Application of Mathematical Signal Processing Techniques to Mission Systems, Research and Technology Organization*, p. 1–31. 66
- A. Khan, L. Sun and E. Ifeachor. 2010, «Learning Models for Video Quality Prediction over Wireless Local Area Network and Universal Mobile Telecommunication System Networks», *IET Communications*, vol. 4, n° 12, p. pages 1389–1403, ISSN 1751-8628. 62
- A. Martin, A. L. Jousselme and C. Osswald. 2008, «Conflict Measure for the Discounting Operation on Belief Functions», dans 11th *International Conference on Information Fusion*, p. 1–8. 65
- A. P. Dempster. 325-339. February 1967, «Upper and Lower Probabilities induced by a Multivalued Mapping», *Annals of Mathematical Statistics Journal*, vol. Volume 38, n° 2. 65
- Adami, D., G. Antichi, R. G. Garroppo, S. Giordano et A. W. Moore. 2015, «Towards an sdn network control application for differentiated traffic routing», dans *2015 IEEE International Conference on Communications (ICC)*, IEEE, p. 5827–5832. 52, 53
- Alreshoodi, M. et J. Woods. June 2014, «Survey on QoE\QoS Correlation Models For Multimedia Services», *Computing Research Repository (CoRR) Journal*, vol. Volume 1306.0221, p. pages 53–72. 61, 67
- Amour, L., S. Souihi et A. Mellouk. October 2017, «Dataset 01 : Controlled Laboratory dataset collecting You Tube QoE IFs using VLC video player. Access : », <https://github.com/Lamyne/QoE-Dataset-VLC-Indicators/>. 63
- Aroussi, S. et A. Mellouk. 2014, «Survey on Machine Learning-based QoE-QoS correlation models», dans *International Conference on Computing, Management and Telecommunications (ComManTel)*, p. 200–204. 61, 67
- Aydin, B., A. A. Kindiroglu, O. Aran et L. Akarun. 2016, «Automatic Personality Prediction from Audiovisual Data using Random Forest Regression», dans *2016 23rd International Conference on Pattern Recognition (ICPR)*, p. 37–42. 62

- P. Le Callet, S. M. et A. Perkis. March 2013, «Qualinet White Paper on Definitions of Quality of Experience (QoE) and Related Concepts», dans *European Network on Quality of Experience in Multimedia Systems and Services*, p. 1–25. 61
- Chen, H., X. Yu et L. Xie. May 2013, «End-to-end Quality Adaptation Scheme based on QoE Prediction for Video Streaming Service in LTE networks», dans *2013 11th International Symposium and Workshops on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt)*, p. 627–633. 62
- CholletKeras, F. 2016, «loss function», <https://keras.io/losses/>. 57
- C.K. Murphy. July 2000, «Combining Belief Functions when Evidence Conflicts», *Decision Support System Journal*, vol. Volume 29, n° 1, p. pages 1–9. 66
- E. Lefevre. June 2002, «Belief function combination and conflict management», *Information Fusion Journal*, vol. Volume 3(2), p. pages 149–162. 66
- Fan, Z. et R. Liu. 2017, «Investigation of machine learning based network traffic classification», dans *2017 International Symposium on Wireless Communication Systems (ISWCS)*, IEEE, p. 1–6. 53
- Fathi-Kazerooni, S., Y. Kaymak et R. Rojas-Cessa. 2019, «Tracking user application activity by using machine learning techniques on network traffic», dans *2019 International Conference on Artificial Intelligence in Information and Communication (ICAIIIC)*, IEEE, p. 405–410. 53
- Florea, M. C., J. Dezert, P. Valin, F. Smarandache et A. Jousselme. April 2006, «Adaptative Combination Rule and Proportional Conflict Redistribution Rule for Information Fusion», *Computing Research Repository (CoRR) Journal*, p. 1–8. 66
- Freund, Y. et R. E. Schapire. pages 148–156. July 1996, «Experiments with a new boosting algorithm», dans *Proceedings of the Thirteenth International Conference On Machine Learning*, Morgan Kaufmann Publisher. 62
- G. Shafer. Princeton University Press, pages 1-314. April 1976, *A Mathematical Theory of Evidence*. 64, 66
- Goodfellow, I., Y. Bengio et A. Courville. 2016, *Deep Learning*, MIT Press. <http://www.deeplearningbook.org>. 54
- H. Mark and F. Eibe and H. Geoffrey and P. Bernhard and R. Peter and W. Ian. June 2009, «The WEKA Data Mining Software : An Update», *ACM SIGKDD explorations newsletter*, , n° 1, p. pages 10–18. 61
- He, K. et J. Sun. 2015, «Convolutional neural networks at constrained time cost», dans *Proceedings of the IEEE conference on computer vision and pattern recognition*, p. 5353–5360. 56
- Ickin, S., K. Wac, M. Fiedler, L. Janowski, J. H. Hong et A. K. Dey. 2012, «Factors Influencing Quality of Experience of commonly used Mobile Applications», *IEEE Communications Magazine*, vol. 50, n° 4, doi :10.1109/MCOM.2012.6178833, p. 48–56, ISSN 0163-6804. 61
- J. Mendes-Moreira, J. C. M. A., Soares et J. Sousa. 2012, «Ensemble Approaches for Regression : A Survey», *ACM Comput. Surv.*, vol. 45, n° 1, p. 10 :1–10 :40, ISSN 0360-0300. 62
- Khan, A., L. Sun et E. Ifeakor. 2009, «Content-Based Video Quality Prediction for MPEG4 Video Streaming over Wireless Networks», *Journal of Multimedia*, vol. 4, n° 4, p. pages 228–239. 62

- L. Amour, S. Souihi, M. S. Mushtaq, S. Hoceini and A. Mellouk. 2016, «Perceived video quality evaluation based on interactive/repulsive relation between the QoE IFs», dans *2016 IEEE International Conference on Communications (ICC)*, p. 1–7. 62
- Langley, A., A. Riddoch, A. Wilk, A. Vicente, C. Krasic, D. Zhang, F. Yang, F. Kouranov, I. Swett, J. Iyengar et collab.. 2017, «The quic transport protocol : Design and internet-scale deployment», dans *Proceedings of the Conference of the ACM Special Interest Group on Data Communication*, ACM, p. 183–196. 52
- Liotou, E., D. Tsolkas et N. Passas. 2016, «A Roadmap on QoE Metrics and Models», dans *2016 23rd International Conference on Telecommunications (ICT)*, p. 1–5. 61
- Lopez-Martin, M., B. Carro, A. Sanchez-Esguevillas et J. Lloret. 2017, «Network traffic classifier with convolutional and recurrent neural networks for internet of things», *IEEE Access*, vol. 5, p. 18 042–18 050. 52, 53, 54
- Lotfollahi, M., M. J. Siavoshani, R. S. H. Zade et M. Saberian. 2020, «Deep packet : A novel approach for encrypted traffic classification using deep learning», *Soft Computing*, vol. 24, n° 3, p. 1999–2012. 53
- M. Galar and A. Fernandez and E. Barrenechea and H. Bustince and F. Herrera. 2012, «A Review on Ensembles for the Class Imbalance Problem : Bagging, Boosting, and Hybrid-Based Approaches», *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 42, n° 4, p. pages 463–484. 62
- Mushtaq, M. S., S. Fowler, B. Augustin et A. Mellouk. 2016, «QoE in 5G Cloud Networks using Multimedia services», dans *IEEE Wireless Communications and Networking Conference*, p. 1–6. 61
- P. K. Mallapragada and R. Jin and A. K. Jain and Y. Liu. 2009, «SemiBoost : Boosting for Semi-Supervised Learning», *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, n° 11, p. pages 2000–2014. 62
- P. Walley. pages 1-700. December 1991, *Statistical Reasoning with Imprecise Probabilities*, Chapman & Hall/CRC Monographs on Statistics & Applied Probability, Taylor & Francis Publisher, ISBN 9780412286605. 65
- Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg et collab.. 2011, «Scikit-learn : Machine learning in python», *Journal of machine learning research*, vol. 12, n° Oct, p. 2825–2830. 57
- Reichl, P., S. Egger, R. Schatz et A. D'Alconzo. 2010, «The Logarithmic Nature of QoE and the Role of the Weber-Fechner Law in QoE Assessment», dans *2010 IEEE International Conference on Communications*, ISSN 1550-3607, p. 1–5. 61
- S. Le Hegarat-Masclé, I. Bloch and D. Vidal-Madjar. 1997, «Application of Dempster-Shafer evidence theory to unsupervised classification in multisource remote sensing», *IEEE Transactions on Geoscience and Remote Sensing*, vol. 35, n° 4, p. pages 1018–1031. 65
- Salman, O., I. H. Elhajj, A. Chehab et A. Kayssi. 2018, «A multi-level internet traffic classifier using deep learning», dans *2018 9th International Conference on the Network of the Future (NOF)*, IEEE, p. 68–75. 53
- Smets, P. 1998, *The Transferable Belief Model for Quantified Belief Representation*, Springer Netherlands, Dordrecht, ISBN 978-94-017-1735-9, p. 267–301, doi :10.1007/978-94-017-1735-9_9. URL https://doi.org/10.1007/978-94-017-1735-9_9. 64, 65, 66

- Smets, P. August 1993, «Belief Functions : The Disjunctive Rule of Combination and the Generalized Bayesian Theorem», *International Journal of Approximate Reasoning*, vol. Volume 9, n° 1, doi :[http://dx.doi.org/10.1016/0888-613X\(93\)90005-X](http://dx.doi.org/10.1016/0888-613X(93)90005-X), p. pages 1–35, ISSN 0888-613X. URL [//www.sciencedirect.com/science/article/pii/0888613X9390005X](http://www.sciencedirect.com/science/article/pii/0888613X9390005X). 66
- Sokolova, M. et G. Lapalme. 2009, «A systematic analysis of performance measures for classification tasks», *Information Processing & Management*, vol. 45, n° 4, p. 427–437. 56
- T. Denoeux. 1995, «A k-Nearest Neighbor Classification Rule based on Dempster-Shafer Theory», *IEEE Transactions on Systems, Man, and Cybernetics*, vol. Volume 25, n° 5, p. pages 804–813, ISSN 0018-9472. 66
- Tantithamthavorn, C., S. McIntosh, A. E. Hassan et K. Matsumoto. 2016, «Automated Parameter Optimization of Classification Techniques for Defect Prediction Models», dans *2016 IEEE/ACM 38th International Conference on Software Engineering (ICSE)*, p. 321–332. 62
- Ujjan, R. M. A., Z. Pervez, K. Dahal, A. K. Bashir, R. Mumtaz et J. González. 2020, «Towards sflow and adaptive polling sampling for deep learning based ddos detection in sdn», *Future Generation Computer Systems*, vol. 111, p. 763–779. 54
- Wang, P., F. Ye, X. Chen et Y. Qian. 2018, «Datanet : Deep learning based encrypted network traffic classification in sdn home gateway», *IEEE Access*, vol. 6, p. 55 380–55 391. 53
- Wang, W., M. Zhu, J. Wang, X. Zeng et Z. Yang. 2017, «End-to-end encrypted traffic classification with one-dimensional convolution neural networks», dans *2017 IEEE International Conference on Intelligence and Security Informatics (ISI)*, IEEE, p. 43–48. 53
- Youssef, Y. B., A. Mellouk, M. Afif et S. Tabbane. December 2016, «Video Quality Assessment Based on Statistical Selection Approach for QoE Factors Dependency», dans *2016 IEEE Global Communications Conference (GLOBECOM)*, p. 1–6. 62
- Zhou, K., W. Wang, C. Wu et T. Hu. 2020, «Practical evaluation of encrypted traffic classification based on a combined method of entropy estimation and neural networks», *ETRI Journal*, vol. 42, n° 3, p. 311–323. 54

Chapitre 5

Contrôle adaptatif du réseau : Le Réseau de demain, un réseau autonome

« Automation applied to an inefficient operation will magnify the inefficiency »

Bill Gates

Resumé

Ce chapitre a pour ambition de présenter **ma vision d'un contrôle adaptatif du réseau**. Ce contrôle adaptatif, qui est en quelque sorte la finalité d'un réseau intelligent, peut s'opérer dans différentes strates du réseau. Les travaux présentés ici se concentrent sur 2 applications en particuliers : le routage et la diffusion des flux vidéo.

Sommaire

5.1 Introduction et contexte	76
5.2 Problématique : qu'est ce qu'un réseau autonome et comment le mettre en oeuvre	76
5.3 Contribution 1 : Routage adaptatif ou comment auto-corriger les défaillances	77
5.3.1 Travaux connexes	78
5.4 Mécanisme adaptatif de routage de segments	79
5.4.1 Compromis exploration-exploitation	80
5.5 Résultats expérimentaux	81
5.5.1 Analyse des performances	82
5.6 Contribution 2 : La diffusion vidéo à débit variable (ABR) sensible à la QoE : Un exemple auto-optimisation	83
5.6.1 Module d'estimation de la qualité de l'image	84
5.7 Évaluation	88
5.8 Synthèses	90
5.8.1 Difficultés rencontrées	90
5.8.2 Forces et faiblesses	91
5.8.3 Possible extension	91
5.8.4 Impact potentiel des travaux	91
5.8.5 Retour d'expérience	91

5.1 Introduction et contexte

5.2 Problématique : qu'est ce qu'un réseau autonome et comment le mettre en oeuvre

Un réseau autonome est un réseau qui offre des capacités "d'auto-*" sans aucune intervention humaine. D'un point de vue architectural, il est basé sur un ensemble d'agents caractérisés par le fait de :

- **se connaître soi-même (self-awareness)** soit la capacité d'un système de se connaître soi-même indépendamment de l'environnement et des autres individus. D'un point de vue philosophique, la connaissance de soi-même peut être résumée par la citation : « cogito ergo sum » (Je pense, donc je suis). D'un point de vue purement technique, un système doit donc avoir connaissance de ses propres ressources, ses composants, ses performances et son état interne.
- **connaître son environnement (environment-awareness)** soit la capacité d'un système à connaître l'ensemble des mécanismes avec lesquels il est en interaction. Cette caractéristique a une incidence sur la faculté du système à s'adapter aux différents contextes.
- **s'auto-monitorer (self-monitoring)** soit la faculté d'un système à suivre par soi-même et régulièrement son comportement et son évolution en vue de pouvoir évoluer et s'améliorer en fonction du contexte.
- **s'auto-ajuster (self-adjusting)** soit la capacité d'un système à s'adapter au changement de son environnement et à rétablir ses fonctionnalités après perturbation, sans aucune intervention externe.

Ces caractéristiques constituent les conditions nécessaires pour que le réseau autonome puisse assurer les 4 objectifs décrits par [Sterritt, 2004] :

- **l'auto-configuration (self-configuring)**. Elle couvre le besoin de rendre la configuration et la reconfiguration d'un réseau et de ses entités dynamiques et autonomes, en particulier, dans les phases d'installation et de suppression d'un nouveau composant dans le réseau. Ce dernier pourra ainsi auto-déterminer la bonne configuration à adopter et éviter les nombreuses erreurs de configuration liées au facteur humain.
- **l'auto-réparation (self-healing)**. Elle permet aux réseaux d'automatiquement identifier, analyser, résoudre et réparer les problèmes. L'objectif étant de maintenir de façon continue la disponibilité des services et de gagner du temps par rapport aux approches actuelles où parfois un certain temps est requis pour le traitement du problème.
- **l'auto-optimisation (self-optimising)**. Elle permet d'automatiser la gestion des performances des réseaux. Ces dernières doivent être optimisées et évaluées de façon continue. En effet, l'optimisation d'un réseau est une tâche difficile à effectuer compte tenu du nombre de paramètres en jeu. Afin d'éviter à l'administrateur de paramétrer lui-même les équipements et les réseaux pour fonctionner correctement, les systèmes doivent continuellement chercher à améliorer leurs performances et optimiser leurs opérations.
- **l'auto-protection (self-protecting)**. Elle concerne les capacités d'un système à se protéger et à se sécuriser. L'automatisation de la sécurité a pour but de développer un comportement pouvant détecter les situations où la stabilité des réseaux et des services peut être remise en cause par le biais d'actions volontaires ou non et de se protéger en conséquence pour ne pas perturber les usagers.

À mon sens, la proposition qui a définitivement ancré le besoin d'une architecture autonome basée sur l'intelligence artificielle est "**A Knowledge Plane for the internet**" [Clark et collab., 2003]. En effet, l'introduction de la notion du "Plan de Connaissance" ou "Knowledge Plane (KP)" a constitué un tournant car elle offre la possibilité à tout réseau, aussi large soit-il, de se connaître soi

même ainsi que tous les évènements qui l'affectent. Cela a constitué un changement fondamental de la conception du réseau de communication en évoluant d'un paradigme "centré réseau", à un paradigme "centré connaissance", ce qui a permis de réfléchir à de nouvelles approches d'auto-monitoring et d'auto-ajustement offrant ainsi de nombreux avantages tels que l'automatisation de la gestion, d'optimisation et de dépannage des réseaux. Néanmoins, peu de prototypes mettant en œuvre cette proposition ont été déployés. La cause principale de cela fut sans doute le manque de programmabilité des équipements. L'émergence de la virtualisation des réseaux et du paradigme SDN (Software-Defined Networking) qui découplent le plan de contrôle du plan de données a levé ce verrou [Feamster et collab., 2014]. En effet, le SDN, par une centralisation du contrôle basée sur une vue globale et unique du réseau, permet de programmer dynamiquement et facilement les différents éléments pour répondre à un certain contexte. Ce chapitre illustre nos modestes contributions dans la mise en place d'une certaine autonomie dans les services réseau.

5.3 Contribution 1 : Routage adaptatif ou comment auto-corriger les défaillances

Les opérateurs de réseau ont été confrontés à un défi lié à l'amélioration des services réseau [Li et collab., 2020]. Concrètement, le développement rapide de la 4G puis de la 5G a amené de nombreuses évolutions en matière de réseau (par exemple, faible latence, haute fiabilité, etc.). Par conséquent, les utilisateurs finaux s'attendent à un traitement différencié des applications. Cependant, il est de plus en plus difficile pour les opérateurs de réseau de mettre en œuvre des politiques de routage correspondant aux exigences des applications en utilisant les moyens usuels qui nécessitent l'intervention humaine. C'est pourquoi de nombreux opérateurs de réseau (par exemple, NTT, Vodafone, etc.) ont, dans un premier temps, mis en œuvre de nouveaux mécanismes de routage de segments (SR : *Segment Routing*) [cisco, January 2021b,J]. L'idée centrale des mécanismes SR est basée sur la notion de routage à la source [Kushwaha et collab., 2020] et de tunneling pour garantir la propriété d'évolutivité en diminuant la quantité d'informations d'état à traiter dans le réseau central. En outre, le principal avantage de SR est de résoudre les problèmes d'évolutivité et les limitations de l'approche MPLS. Concrètement, le SR nécessite peu de maintenance d'état dans les nœuds du réseau central. En outre, il tire parti du routage ECMP et de la mise en œuvre sur des AS multiples [Kukreja et collab., 2016]. Cette technique a rendu possible les approches de routage adapté aux applications [Kukreja et collab., 2016; Peng et collab., 2020] en bénéficiant des avancées permises par la mise en réseau définie par logiciel (SDN). Cependant, les implémentations existantes de routage de segments présentent deux inconvénients majeurs :

- Tout d'abord, ce mécanisme de routage considère la classe d'application pour répondre à différentes exigences SLA (Service-level Agreement), mais l'identification de l'application est parfois compliquée en raison du trafic réseau qui est de plus en plus chiffré. Dans le passé, la classe d'application pouvait être obtenue par le DSCP (Differentiated Services Code Point), donc, directement depuis l'en-tête du paquet ou par une inspection approfondie des paquets eux-mêmes. Cependant, de nos jours, le champ DSCP est souvent modifié pendant la transmission du paquet [Custura et collab., 2018] tandis que l'inspection approfondie des paquets n'est pas efficace dans un contexte de trafic chiffré [Dong et collab., 2020]. En effet, selon un récent rapport de CISCO [February 2021], en 2019, 80% du trafic web était chiffré, contre 40% en 2016. Par conséquent, il est nécessaire d'adopter une nouvelle approche de classification du trafic pour identifier la classe d'application à l'image de ce qui a été décrit dans le précédent chapitre.
- Deuxièmement, l'identification des chemins de routage se base toujours sur un certain degré d'expertise humaine, ce qui ne permet pas de s'adapter aux changements réseau. D'autant plus qu'il est nécessaire, de nos jours, d'assurer un certain niveau de qualité de l'expérience (QoE) pour garantir les exigences SLA de l'utilisateur.

Dans cette contribution, nous avons donc proposé une nouvelle approche autonome qui consiste en un mécanisme adaptatif de routage de segments basé sur un réseau SDN. Ce système est robuste, autonome et adapté au chiffrement du trafic réseau.

Dans la littérature, il existe un certain nombre d'approches de routage de segments qui tiennent compte de la QoE. Cependant, ils ne sont pas efficaces avec les réseaux dynamiques. Or, l'un des aspects qui a le plus motivé mes travaux tout au long de ces années est la prise en compte de cette dynamique, particulièrement, dans les réseaux de grande taille. Nous avons donc proposé une nouvelle méthode, basée principalement sur une technique d'apprentissage par renforcement [Tong et collab., 2020].

5.3.1 Travaux connexes

Cette section présente les travaux connexes à cette contribution : le routage adapté aux applications et les méthodes de classification du trafic crypté.

Barakabitze et collab. [2020] ont proposé QoEMultiSDN, un algorithme de routage source multi-chemins. Concrètement, cette approche sélectionne les chemins de routage en utilisant un modèle de plus court chemin. Bien que QoEMultiSDN sélectionne de manière adaptative le débit binaire de la diffusion vidéo en utilisant la qualité d'expérience, le chemin de routage est toujours identifié en fonction des paramètres de qualité de service.

Li et collab. [2020] ont souligné que les opérateurs de réseau ont été confrontés au défi de fournir de meilleurs services avec la croissance rapide de la 5G et des services multimédias qui nécessitent diverses exigences de réseau (par exemple, faible latence, haute fiabilité, etc.). Cependant, les opérateurs de réseau ne savent pas quelles sont les applications présentes dans leurs réseaux, ils sont donc incapables d'obtenir une vue globale pour gérer efficacement les réseaux. Par conséquent, de nombreux travaux de recherche se concentrent sur le routage adapté aux applications (par exemple, MPLS adapté aux applications, etc.).

Rego et collab. [2019] ont proposé une amélioration du protocole de routage Open Shortest Path First (OSPF) dans les environnements SDN. Cette approche modifie le calcul des métriques de manière adaptative en fonction des différentes applications afin de sélectionner les chemins de routage appropriés. Cependant, l'identification d'un calcul métrique est parfois compliquée. De même, **Bahnasse et collab. [2018]** ont proposé un mécanisme SDN basé sur le protocole MPLS, sensible aux applications pour optimiser les ressources du réseau. Cette approche identifie le trafic en fonction du champ DSCP de l'en-tête du paquet. Ensuite, une politique de routage spécifique est appliquée pour chaque application afin de respecter les contraintes de bande passante. Cependant, l'identification de la classe d'application à l'aide du champ DSCP est parfois inexacte car ce dernier peut être modifié pendant la transmission des données [Custura et collab., 2018]. Google [Yap et collab., 2017] a proposé Espresso, une infrastructure de routage pour le peering Internet basée sur le SDN qui offre une opportunité pour mettre en oeuvre un mécanisme basé sur le protocole MPLS et sensible aux applications. Espresso a permis d'augmenter de 13 % le trafic réseau sur ses infrastructures et d'améliorer l'utilisation des liens et, par conséquent, la perception des utilisateurs par rapport au routage basé sur le protocole BGP (Border Gateway Protocol). Néanmoins, Espresso souffre de plusieurs obstacles principalement liés à l'extensibilité.

Kukreja et collab. [2016] ont présenté une démonstration de routage de segments basé sur SDN pour les réseaux multi-domaines. Dans cette démonstration, un orchestrateur trouve les chemins de routage appropriés. L'objectif est de répondre aux différentes exigences de ces applications en matière de ressources (par exemple bande passante, contraintes de délai, etc.). Néanmoins, l'approche nécessite que la classe de l'application soit connue par l'orchestrateur. **Peng et collab. [2020]** ont proposé une approche sensible aux applications qui tire également parti du mécanisme de routage de segments afin de garantir un certains niveau de SLA. Cette approche identifie les caractéristiques de l'application via un mécanisme d'inspection approfondie des paquets, puis achemine les paquets vers les chemins correspondants (politique ou tunnel d'ingénierie du trafic). Néanmoins, l'affectation des chemins repose en partie sur une intervention humaine, ce qui n'est

pas efficace dans un environnement réseau dynamique.

5.4 Mécanisme adaptatif de routage de segments

La figure 5.1 représente un système modulaire que nous avons proposé afin de mettre en place un mécanisme de contrôle adaptatif dans un réseau (dans notre exemple, un routage de segments).

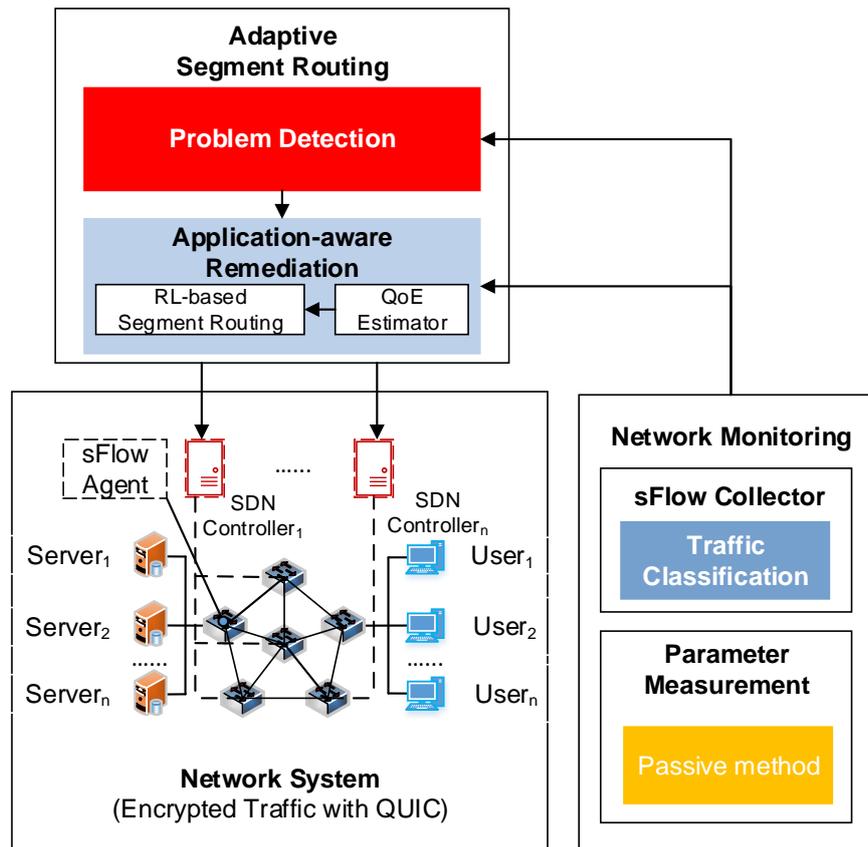


FIGURE 5.1 – Le cadre du mécanisme SR adaptatif basé sur le SDN.

À l'initialisation, le flux réseau est transmis en utilisant un algorithme de routage de segment (SR) classique basé sur l'approche des plus courts chemins standard. Ensuite, les paramètres du réseau provenant du module de mesures sont analysés dans le module de détection des anomalies pour détecter tout comportement suspect (par exemple, l'augmentation de la latence, la perte de paquets, etc.) afin de déclencher le mécanisme adaptatif. Le mécanisme proposé considère la classe d'application pour mettre en œuvre des politiques de routage appropriées correspondant à chaque nature de flux. Lorsque le trafic réseau est chiffré, un module de classification du trafic permet d'identifier cette information.

L'algorithme SR mise en oeuvre est formalisé comme une tâche d'apprentissage par renforcement (RL : Reinforcement Learning) qui contient *un agent*, *un état*, *une action*, *une récompense* et *une politique* (Fig. 5.2). Le détail est décrit comme suit :

Agent : Une entité du système qui applique un algorithme d'apprentissage pour effectuer ses tâches. Dans un problème de routage, un agent sélectionne les chemins appropriés pour optimiser une récompense.

État (s) : Un instantané de l'environnement du réseau qui est observé par l'agent.

Action (a) : Une action qui illustre la façon dont un agent répond à l'environnement réseau. Dans le problème du routage, l'action est un chemin de routage entre un serveur et un client dans

le réseau. Tous les chemins de routage du réseau peuvent être obtenus dans le contrôleur SDN.

Politique (π) : Une politique est une correspondance entre un état observé et une action dans l'environnement réseau.

Récompense (r) : Une récompense est un retour de l'environnement à l'agent. Dans le problème de routage, l'agent surveille un état du réseau s et exécute une action a de la politique de routage. Ensuite, l'agent passe à l'état suivant s et reçoit une récompense r . La récompense est le score MOS d'un chemin choisi qui est calculé via un estimateur de qualité d'expérience.

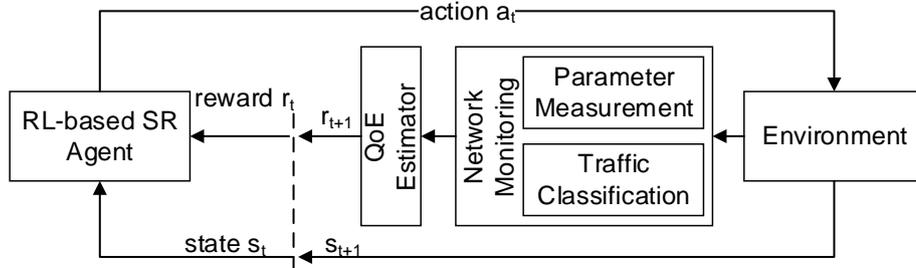


FIGURE 5.2 – Le mécanisme SR basé sur RL

L'objectif principal d'une approche d'apprentissage par renforcement est d'optimiser une fonction objectif O_f (Eq. 5.1) :

$$O_f = \text{Max} E \left[\sum_{t=0}^{\infty} \gamma_t \times r_t \right]. \quad (5.1)$$

où $\gamma_t \in [0, 1]$ est un facteur d'actualisation.

Il existe deux types d'approches d'apprentissage par renforcement : les approches basées sur un modèle et les approches sans modèle. Dans les premières, les agents apprennent le modèle d'environnement et améliorent leurs politiques pour obtenir l'optimalité, tandis que dans les secondes, les agents optimisent leurs politiques sans information préalable sur l'environnement. Les premières approches permettent d'apprendre plus rapidement que les secondes. Elles sont néanmoins moins populaires en raison du coût important et de la dépendance à l'égard de la précision d'une information initiale : [Mammeri, 2019]. Par conséquent, l'approche sans modèle est utilisée dans ce travail. Dans cette approche, la valeur Q estime la valeur d'une action donnée dans un état donné. $Q(s, a)$ est le rendement attendu en partant de l'état s et en effectuant l'action a suivant la politique π . À un pas de temps donné, l'agent dans l'état s , effectue l'action a , reçoit la récompense r et passe à l'état suivant s' . La valeur Q est mise à jour comme dans l'équation 5.2 :

$$Q(s, a) = (1 - \alpha)Q(s, a) + \alpha[r + \gamma \max_{a'} Q(s', a')]. \quad (5.2)$$

où α est un taux d'apprentissage et γ est un facteur d'actualisation.

La complexité temporelle de l'algorithme d'apprentissage par renforcement est de $O(m * N_0)$ où N_0 est le nombre d'actions, et m est la taille de l'espace d'états [[Koenig et Simmons, 1993; Jin et collab., 2020].

5.4.1 Compromis exploration-exploitation

Dans un processus d'apprentissage par renforcement, les phases d'exploration et d'exploitation doivent être équilibrées pour obtenir un score MOS cumulatif optimal (récompense cumulative). La phase d'exploitation, qui sélectionne le chemin de routage (action) avec la valeur Q maximale, ne peut pas être mise en œuvre systématiquement car chaque chemin de routage doit être évalué fréquemment pour obtenir le score MOS optimal. Dans ce travail, le compromis entre les phases d'exploration et d'exploitation est formalisé sous la forme d'un problème MAB (Multi-Armed Bandit). Le problème MAB est une formalisation de tâches décisionnelles séquentielles. À un moment donné, un décideur choisit une action et reçoit une récompense provenant

d'une distribution inconnue correspondant à cette action. L'objectif principal est de maximiser la récompense totale reçue par une séquence d'actions. Nous avons comparé trois algorithmes de sélection ϵ -greedy [Temlyakov, 2015], softmax [He et collab., 2018], et UCB1 (Upper Confidence Bounds) [Tran et collab., 2013].

L'algorithme ϵ -greedy ou *epsilon-greedy* est le plus simple pour résoudre le problème du bandit multi-armé. Concrètement, l'agent sélectionne le chemin de routage avec la valeur Q la plus élevée avec une probabilité de $(1-\epsilon)$. Sinon, l'agent sélectionne le chemin de routage de manière aléatoire. Ensuite, la valeur ϵ peut diminuer au fil du temps, au fur et à mesure que l'agent gagne en confiance.

L'algorithme softmax sélectionne les chemins de routage en fonction d'une fonction de probabilité de la valeur Q. Chaque chemin de routage a_i est affecté à une probabilité p_i comme décrit dans l'équation 5.3 :

$$p_i = \frac{e^{\frac{Q_{a_i}}{\tau}}}{\sum_{j=1}^{N_0} e^{\frac{Q_{a_j}}{\tau}}} \quad (5.3)$$

où τ est un paramètre de température, N_0 est un nombre de chemins de routage et Q_{a_j} est une valeur Q du chemin de routage a_j .

Lorsque le paramètre de température τ est réduit, les chemins de routage sont exploités plus fréquemment. Ainsi, on peut par exemple réduire le paramètre de température τ à chaque épisode (temps d'acheminement) afin de favoriser l'exploration des autres chemins. Par conséquent, l'algorithme softmax explore non seulement les chemins de routage les moins utilisés mais garantit également la sélection du meilleur chemin de routage.

Enfin, l'algorithme UCB1 est basé sur un indice UCB qui est défini comme la somme d'une valeur Q actuelle et d'une limite de confiance. L'indice UCB est décrit par l'équation 5.4 :

$$UCB - index_{a_i} = Q_{a_i} + \sqrt{\frac{2 \ln(N)}{n_{a_i}}}. \quad (5.4)$$

où Q_{a_i} est une valeur Q du chemin de routage a_i , n_{a_i} est un nombre de temps choisis du chemin de routage a_i et N est un nombre d'épisodes (temps d'acheminement).

Après avoir calculé l'indice UCB pour chaque chemin de routage, l'algorithme UCB1 sélectionne le chemin avec l'indice UCB maximal. Comme le montre l'équation 5.4, l'indice UCB comprend deux parties, dont la valeur Q_{a_i} et la limite de confiance $\sqrt{\frac{2 \ln(N)}{n_{a_i}}}$. Un chemin d'acheminement est choisi lorsque la valeur de Q est grande ou lorsque la limite de confiance est élevée. Lorsque le chemin de routage avec la grande valeur Q est choisi, ce choix est un essai d'exploitation. Lorsque la limite de confiance est élevée, ce choix est un essai d'exploration. La limite de confiance est plus élevée lorsque le nombre de moments choisis du chemin de routage est plus petit par rapport aux autres chemins. En d'autres termes, moins le chemin de routage est choisi, plus il a de chances d'être choisi.

5.5 Résultats expérimentaux

Concernant les expériences, nous avons utilisé l'outil mininet [min, February 2021] pour émuler un réseau avec des commutateurs Openflow. Nous avons implémenté un contrôleur SDN en utilisant ONOS [ono, February 2021] et l'avons connecté aux commutateurs Openflow dans le réseau émulé par mininet. Nous avons considéré 2 topologies Leaf-Spine en variant le nombre de nœuds. Pour générer le trafic du streaming vidéo dans le réseau, nous rejeuons un fichier pcap du trafic QUIC en utilisant l'application tcpreplay. Ce fichier pcap est collecté en regardant des vidéos sur Youtube avec Google Chrome. Pour le compromis exploration-exploitation, les valeurs de ϵ et τ sont respectivement fixées initialement à 1 et 2. ϵ décroît linéairement avec le temps. Pour l'algorithme RL, α et γ sont fixés à 0,7 et 1. Ces paramètres sont choisis en fonction d'un travail existant

[Lin et collab., 2016]. Enfin, dans notre politique de recherche reproductible et ouverte, le code source proposé est disponible sur [Tong, February 2021].

Pour valider les performances du mécanisme SR proposé, notre proposition est comparée aux benchmarks suivants :

1. *SR standard basé sur le SDN (Standard_SR)* : Cet algorithme utilise l'algorithme de Dijkstra pour déterminer les chemins les plus courts entre les serveurs et les clients.
2. *SR basé sur le SDN avec QoE maximale (Max_QoE)* Grigoriou et collab. [2017] : Cette approche exhaustive calcule le score MOS de tous les chemins de routage et sélectionne les chemins avec un score MOS maximal.

Ces approches sont évaluées à l'aide des mesures de performance suivantes :

1. *MOS* : qualité perçue du côté de l'utilisateur.
2. *CPU Usage* : pourcentage de la capacité du CPU qui est calculé via la commande *ps* (état du processus) dans les systèmes d'exploitation de type Unix.
3. *Control Overhead ou surcharge de contrôle* : rapport entre le nombre de paquets de contrôle et le nombre total de paquets envoyés. En effet, pour découvrir la topologie du réseau et mettre à jour l'état des liens (par exemple, latence, perte, etc.), un algorithme de routage doit générer des paquets de contrôle (par exemple, des paquets LLDP, etc.).

5.5.1 Analyse des performances

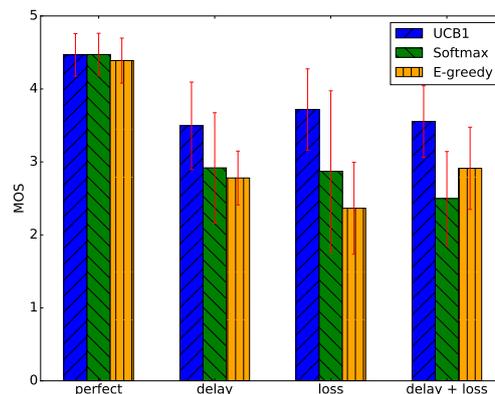


FIGURE 5.3 – Score MOS moyen et écart type des trois algorithmes de sélection.

La figure 5.3 illustre le score MOS moyen et l'écart type de *UCB1*, *softmax* et *ϵ -greedy*. Dans le scénario *parfait*, il n'y a ni retard ni perte dans le réseau. Le débit d'envoi est de 2,5 Mbps alors que la capacité du lien est de 10 Mbps. Par conséquent, la modification des chemins de routage n'entraîne pas la fluctuation du score MOS dans les trois algorithmes de sélection. Bien qu'il n'y ait pas de différence significative entre ces algorithmes, le score MOS de *softmax* est légèrement meilleur que les autres. Le score MOS moyen de *softmax*, *UCB1* et *ϵ -greedy* atteint 4,47, 4,46 et 4,39, respectivement. En outre, leurs écarts types sont de 0,29, 0,291 et 0,31, respectivement. Dans ce scénario, le score MOS varie de 4,3 à 4,5. Le score MOS de *ϵ -greedy* converge vers 4,3 tandis que le score MOS des autres converge vers 4,5. Par conséquent, le score MOS moyen de *ϵ -greedy* est inférieur à celui des deux autres algorithmes de sélection.

Le score MOS moyen de *UCB1* est le plus élevé, et l'écart type est le plus faible parmi les trois algorithmes de sélection pour les autres scénarios. Le score MOS moyen de *UCB1* dans le scénario de retard, le scénario de perte et le scénario final est de 3,55, 3,7 et 3,5, respectivement. *UCB1* explore d'abord le chemin de routage pendant les premiers épisodes et converge ensuite vers le score MOS optimal. Lorsque les états du réseau sont modifiés, *UCB1* peut explorer les chemins

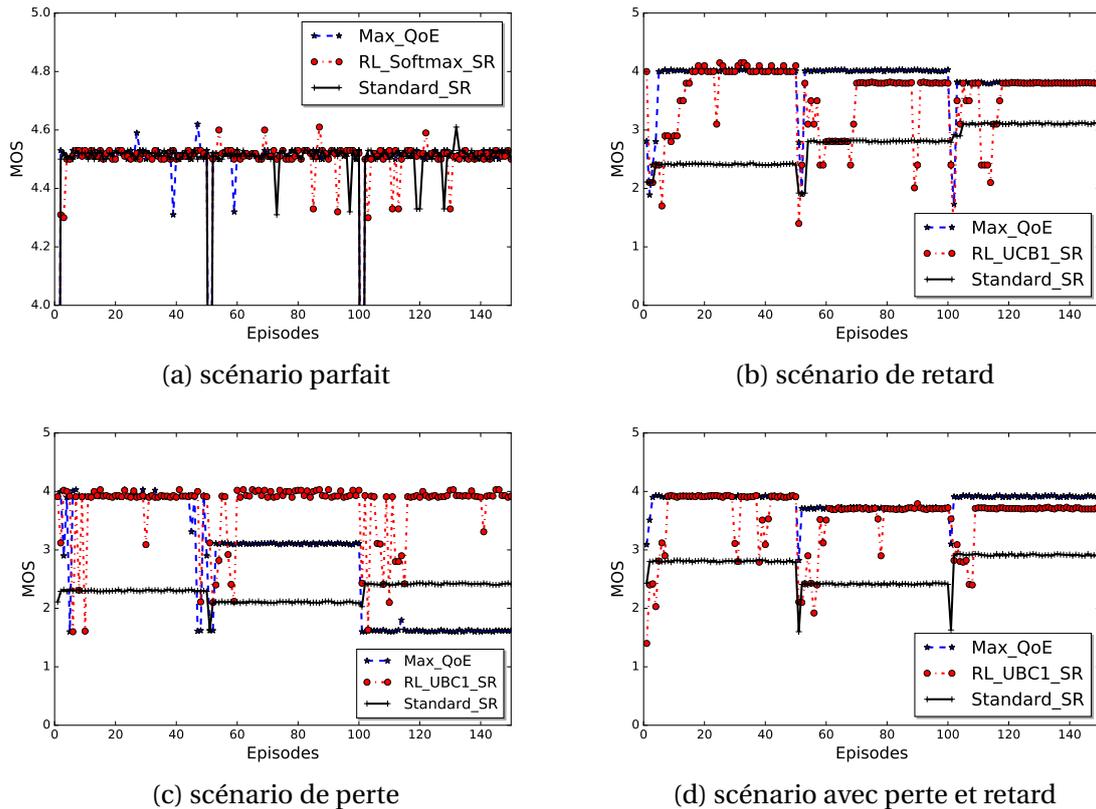


FIGURE 5.4 – Le score MOS de trois mécanismes SR.

de routage les moins utilisés. Par conséquent, il continue à atteindre rapidement une valeur optimale. *Softmax* choisit les chemins de routage en fonction de la fonction de probabilité, il prend donc beaucoup de temps pour converger vers une valeur optimale. Par conséquent, le score MOS de *softmax* est inférieur à celui de *UCB1*. Pour *ϵ -greedy*, il explore d'abord les chemins de routage avec une valeur ϵ élevée. Lorsque la valeur de ϵ se réduit à environ 0, il ne peut plus explorer les chemins de routage fréquemment. Par conséquent, le score MOS de *ϵ -greedy* est inférieur à celui des deux autres algorithmes de sélection.

5.6 Contribution 2 : La diffusion vidéo à débit variable (ABR) sensible à la QoE : Un exemple auto-optimisation

La seconde contribution présentée dans ce chapitre concerne une autre couche du modèle OSI. En effet, on s'est intéressé particulièrement à la vidéo et notamment au mécanisme de streaming. Selon le rapport [CISCO Systems \[June 2017\]](#), le trafic vidéo représente plus de 80% du trafic mondial. Ce trafic est principalement dû aux services de vidéo à la demande fournis par des plateformes telles que YouTube¹, DailyMotion² et Netflix³. Les utilisateurs finaux de ces plateformes émettent une requête HTTP pour demander une séquence vidéo au fournisseur de services. Pour obtenir la qualité appropriée, les fournisseurs de services utilisent des méthodes ABR (Adaptive BitRate Streaming). Dans ces méthodes, la vidéo brute est codée dans différentes représentations stockées sur le serveur web, où la vidéo est divisée en segments de quelques secondes. La structure du segment et ses paramètres associés sont publiés dans un fichier XML appelé "*manifeste*". Le lecteur vidéo télécharge le fichier manifeste et utilise l'une des stratégies ABR pour sélectionner séquentiellement les segments vidéo avec différentes représentations.

1. <https://www.youtube.com/>
 2. <https://www.dailymotion.com/>
 3. <https://www.netflix.com/>

La littérature propose de nombreuses méthodes ABR. Ces méthodes sélectionnent la meilleure qualité de segment vidéo afin d'assurer une bonne expérience utilisateur (QoE). Elles sont classées en trois catégories : les méthodes basées sur le débit (TBA) [Azumi et collab., 2015; J. Jiang et Zhang, January 2014; C. Liu et Gabbouj, February 2011. p. 169-174; Sun et collab., 2016], les méthodes basées sur la mémoire tampon (BBA) [T.Y Huang et Watson, August 2014; Zhou et collab., 2013; V. Krishnamoorthi et Petajan, June 2017; Zahran et collab., May 2016], et des méthodes hybrides [Spiteri et collab., 2016; McLean, 2009; Joseph et de Veciana, 2014; Mushtaq et collab., 2015; Juluri et collab., 2015; Menkovski et Liotta, 2013; M. Claeys et collab., May 2013], utilisant des combinaisons de méthodes TBA et BBA.

Contribution principale

Dans ce contexte, nous avons proposé une nouvelle méthode ABR appelée "*QoE-aware adaptive bit rate (Q2ABR)*". Cette méthode est basée sur l'estimation de la qualité d'expérience de l'utilisateur en termes de score moyen d'opinion (MOS) [ITU, July 2016. p. 1-42] comme entrée du module d'apprentissage par renforcement qui sélectionne la meilleure représentation vidéo. Ceci a pour but d'assurer une QoE correcte pour l'utilisateur et d'optimiser l'utilisation de la bande passante.

La figure (5.5) illustre le schéma général de l'architecture de notre proposition. La méthode

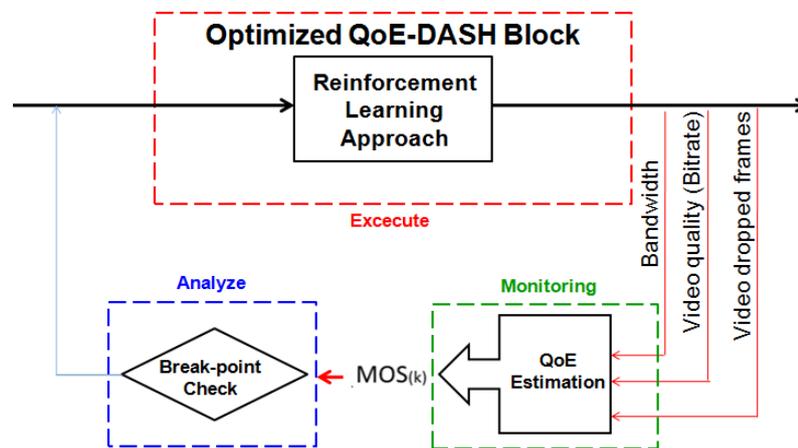


FIGURE 5.5 – Architecture de la méthode de streaming à débit binaire adaptatif (ABR) proposée

Q2ABR se compose de trois modules individuels, l'un pour l'estimation de la qualité d'expérience à l'aide d'une méthode d'apprentissage automatique, l'un pour le contrôle de la variation de cette même qualité d'expérience à l'aide d'un mécanisme de détection des points de rupture, et le dernier pour la sélection de la représentation vidéo à l'aide d'un mécanisme d'apprentissage par renforcement.

5.6.1 Module d'estimation de la qualité de l'image

Comme indiqué dans la littérature, plusieurs méthodes peuvent être utilisées pour évaluer la qualité d'expérience d'un service vidéo. La QoE est une métrique qui peut être évaluée directement par des humains (méthodes dites subjectives) ou estimée automatiquement par des paramètres de réseau et de performance (méthodes dites objectives). Nous pouvons également considérer les méthodes hybrides comme une combinaison des deux méthodes précédentes. Ces méthodes sont basées sur des méthodes d'apprentissage automatique, et elles utilisent un ensemble de données subjectives comme entrée pour le processus d'apprentissage. L'avantage de ces méthodes est l'efficacité et l'exécution en temps réel [Liotou et collab., 2016]. Le premier module se base sur l'estimateur de QoE présenté dans le chapitre précédent. Le deuxième module de la méthode Q2ABR consiste à surveiller la variation de la qualité d'expérience à l'aide du mécanisme de détection des points d'arrêt (BPD) [R. Bertrand et collab., 2010. p 1-26]. Ce module

appelé "BP-QoE" a pour objectif d'observer les changements dans la qualité d'expérience. Autrement dit, le module surveille les changements dans les scores MOS estimés. L'objectif est de déterminer le moment où il convient de lancer l'action appropriée (changement de représentation vidéo) pour assurer une QoE optimale. Ce mécanisme est présenté en termes mathématiques comme suit : considérons une séquence d'observations MOS = (MOS₁, MOS₂, ..., MOS_L) de longueur L que nous calculons périodiquement. Le MOS est distribué indépendamment dans le temps (dynamique du système due à la variation des paramètres impactant la perception de l'utilisateur). Le mécanisme de détection des points de rupture consiste alors à segmenter la séquence MOS, détecter le nombre de points de rupture N, puis à identifier les positions des points de rupture C = (c₁, c₂, ..., c_N), où 1 = c₀ < c₁ < ..., c_{N+1} + L. Ainsi, un point de rupture (c_i) est la limite entre deux segments adjacents. Dans notre cas, c_i représente les positions des points de rupture MOS (MOS_B). Pour détecter ces positions, nous avons utilisé la méthode de l'entropie croisée (EC) [Priyadarshana et Sofronov, 2015]. Cette méthode consiste en une procédure d'optimisation stochastique [RY. Rubinstein, May 1997] qui utilise des distributions de probabilité.

Le troisième module de la méthode Q2ABR consiste en l'implémentation de la méthode de streaming à débit binaire adaptatif. À cette fin, différentes approches peuvent être utilisées. Dans notre cas, nous avons choisi d'utiliser la technique d'apprentissage par renforcement (Figure 5.6).

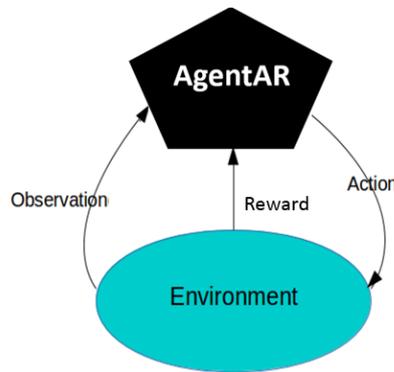


FIGURE 5.6 – Description de l'agentABR

Dans cet environnement, l'"AgentABR" exécute des actions, puis reçoit des changements d'état et des récompenses du système [Menkovski et Liotta, 2013]. En termes mathématiques, le tuple {S, A, T, R, γ} définit les composants du système d'apprentissage par renforcement, où S définit les états, A représente les actions, T indique les fonctions de transition du système, r représente les récompenses du système et γ représente un paramètre de mise à jour qui sera utilisé par la politique de contrôle (noté π). Ce paramètre de mise à jour est une association d'une action a_i avec chaque état s_d qui maximise la récompense moyenne à long terme (r_i), définie comme gain (G) et calculée avec l'équation (5.5) :

$$G = E^{\pi} \left[\sum_{i=1}^{\infty} (\gamma^i r_i) \right] \quad (5.5)$$

Dans le système d'apprentissage par renforcement que nous proposons, nous représentons un état par (s_j ∈ S) comme représentation de segment vidéo (*quality*) et utilisons trois types d'actions :

- (1) augmenter la qualité de la vidéo (**a₁ = qualité + +**),
- (2) réduire la qualité vidéo (**a₂ = quality - -**),
- (3) utiliser la même qualité (invariant) (**a₃ = quality**).

La fonction d'apprentissage du processus de sélection de la représentation vidéo consiste à sélectionner la meilleure action a_i pour passer d'un état (représentation vidéo) à un autre, si nécessaire. Par conséquent, pour sélectionner a_i parmi toutes les actions possibles $A = \{a_1, a_2, a_3\}$, nous définissons ce qui suit :

- Fonction de récompense r comme valeur de la QoE estimée actuelle à l'aide du premier module de la méthode Q2ABR.
- Pénalité γ , qui dépend des variations des valeurs MOS prédites résultant du deuxième module de la méthode Q2ABR (BP-QoE), dans lequel l'entropie croisée [Priyadarshana et Sofronov, 2015] est utilisée pour détecter la position du point de rupture MOS final (MOS_B).

Par conséquent, l'action sélectionnée (a_i) est déduite du calcul de la pénalité (γ_{a_i}) en utilisant (i) les scores moyens de MOS_t ainsi que (ii) la variabilité de ces scores (l'écart type des valeurs de MOS_t). L'équation (5.6) montre le calcul de la pénalité (γ_{a_i}) pour chaque action a_i . Parmi toutes les actions considérées, celle qui présente une pénalité minimale γ_{a_i} a été sélectionnée pour décider de la qualité à utiliser pour le segment vidéo suivant.

$$\gamma_{a_i} = abs(mean(MOS_{previous}) - mean(MOS_{actual})) + abs(mean(MOS_{previous}) - MOS_{a_i}) \quad (5.6)$$

Avec :

- MOS_{a_i} : valeur MOS instantanée estimée avec le premier module a_i .
- $MOS_{previous}$: l'ensemble des précédentes valeurs MOS prédites MOS_B.
- MOS_{actual} : l'ensemble de toutes les valeurs MOS prédites

D'un point de vue fonctionnel, la solution Q2ABR est basée sur la norme DASH [Han et collab., 2016]. L'objectif est d'optimiser en temps réel la représentation vidéo des segments en fonction de la mesure instantanée de la QoE calculée en termes de scores MOS. La figure (5.7) illustre le fonctionnement de ce processus.

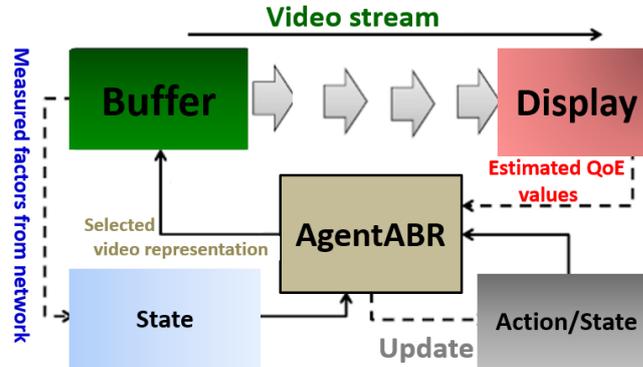


FIGURE 5.7 – Description du système RL

Comme la méthode 'OSCAR' [Zahran et collab., May 2016], la solution Q2ABR considère trois niveaux de tampon. Ceux-ci sont représentés comme suit :

$$0 < faible < \alpha_1 < moyenne < \alpha_2 < élevé. \quad (5.7)$$

Où α_1 et α_2 sont les deux frontières qui délimitent les trois zones et sont calculées expérimentalement.

- Au niveau inférieur, la représentation vidéo la plus basse est nécessaire pour accélérer le remplissage du tampon. En d'autres termes, les deux actions a_2 et a_3 sont privilégiées.

- Au niveau supérieur, le client vise la livraison d’une représentation vidéo élevée tout en évitant les périodes d’arrêt (mise en mémoire tampon vide). En d’autres termes, les deux actions a_1 et a_2 sont préférées.
- Au niveau moyen, Q2ABR teste les trois actions définies précédemment et sélectionne l’action optimale déterminée par le calcul de la pénalité (γ_{a_i}) en utilisant l’équation (5.6).

Le diagramme présenté dans la figure (5.8) résume notre solution.

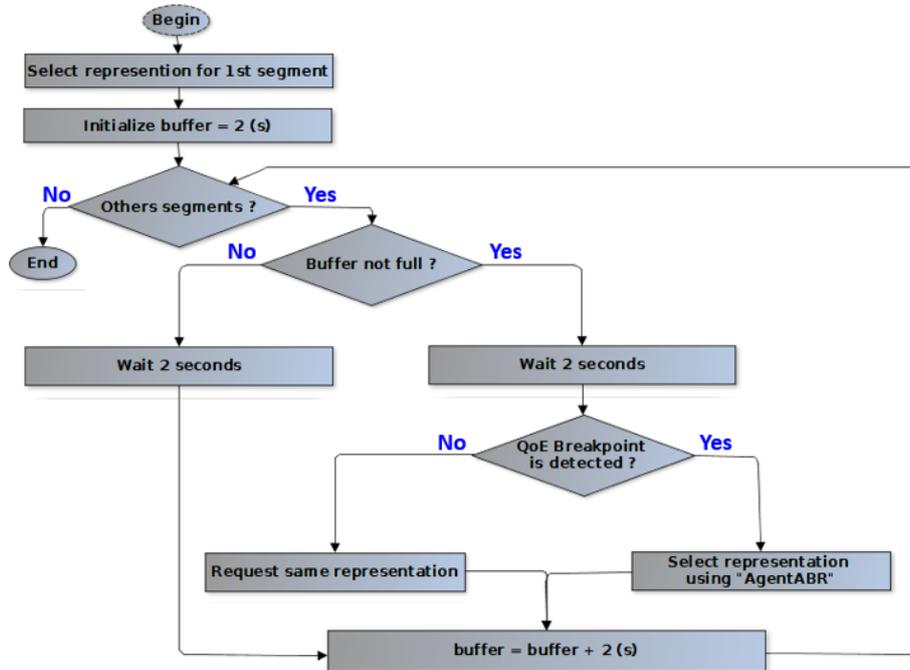


FIGURE 5.8 – Exécution de l’approche Q2ABR

Pour commencer, la représentation vidéo du premier segment est sélectionnée en fonction du débit binaire vidéo le plus proche de la moitié du débit régulier mesuré. Par exemple, si le débit mesuré est égal à 1600 Kb/s, la représentation choisie sera ‘4’ (640x480, 1Mb/s) avec un débit vidéo de 1 Mb/s. Ce choix est justifié par le fait que la solution Q2ABR privilégie les petites représentations vidéo au début afin d’assurer un bon taux de remplissage du tampon et de le remplir rapidement pour éviter un tampon vide. Ensuite, le module d’estimation de la qualité d’expérience calcule périodiquement la qualité d’expérience en fonction du score MOS. Au cours de nos expériences, nous avons opté pour une période égale à la durée d’un segment vidéo, que nous avons fixée à 2 secondes, comme le recommande la norme DASH [Han et collab., 2016]. Lorsque le score MOS est calculé, le module de surveillance de la variation de la qualité d’image BP-QoE évalue l’existence d’un point de changement dans les scores MOS calculés, comme le montre la figure (5.9).

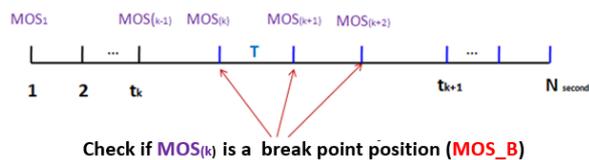


FIGURE 5.9 – Processus de détection du point de rupture

Si aucun point de changement (MOS_B) n’est détecté et que le tampon n’est pas rempli, la même représentation vidéo est demandée pour le segment suivant. Lorsqu’un point de changement (MOS_B) est détecté et que le tampon n’est pas rempli, l’AgentABR initie les actions nécessaires selon la formule (5.6) pour définir la représentation vidéo adéquate pour le segment suivant.

5.7 Évaluation

Les performances de la solution Q2ABR sont évaluées en comparaison à 3 autres solutions ABR existantes : HLS [McLean, 2009], BOLA [Spiteri et collab., 2016] et OSCAR [Zahran et collab., May 2016]. Nous avons fixé la taille de la mémoire tampon à 10 secondes, et nous utilisons un clip vidéo d'une durée de 60 secondes. Ce clip est codé en segments de deux secondes avec différentes représentations (125, 180 x 144), (250, 380 x 240), (1000, 720 x 480), (2000, 1280 x 720), (4000, 1920 x 1080). L'émulation des conditions du réseau est basée sur des traces réelles collectées au cours de trajets effectués par différents modes de transport (bus [scénario 'A'], RER [scénario 'B'], et train [scénario 'C']) dans les villes de Paris, Créteil, et Vitry sur Seine (France). Ces traces sont librement accessibles.

Le nombre de rebuffering résultant des différentes méthodes ABR a un impact direct sur la QoE des utilisateurs, comme le démontrent les études Zhou et collab. [2013]; T.Y Huang et Watson [August 2014]. Ces études expliquent que, si ce nombre est important, il peut conduire au mécontentement des utilisateurs ou à l'abandon du service.

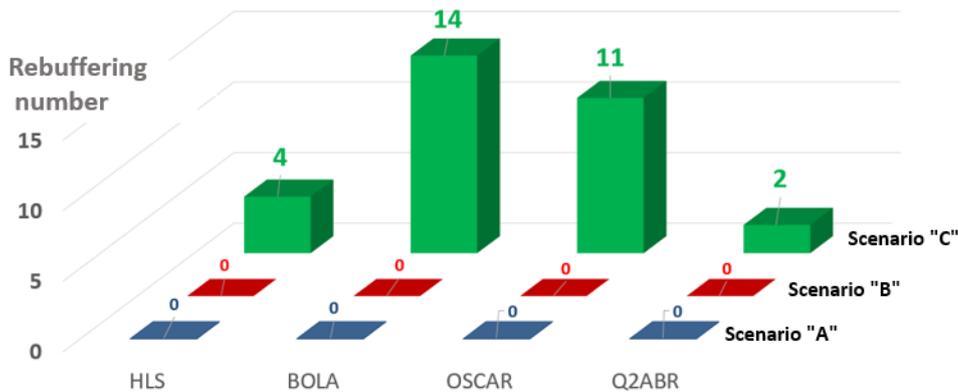


FIGURE 5.10 – Résultats de la comparaison des approches ABR en termes de nombre de rebufferings

À partir de la figure (5.10), nous pouvons observer que la solution Q2ABR évite les situations où le tampon se retrouve vide pour les scénarios A et B. Cela s'explique par les conditions de réseau acceptables. Dans le cas du scénario C, la solution Q2ABR assure un nombre minimum de rebuffering par rapport aux autres ABRs testés.

En ce qui concerne le temps de remplissage du tampon, la solution Q2ABR garantit le meilleur taux de remplissage de la mémoire tampon avec un minimum de 5,6 secondes dans le cas du scénario C (conditions de réseau moins bonnes). Ce résultat justifie le nombre moyen de rebuffering réduit n_{st} que nous avons illustré précédemment.

D'après la figure (5.11), nous pouvons observer que le changement de qualité vidéo (n_{sw}) varie en fonction du scénario. La solution Q2ABR surpasse les autres solutions en terme de nombre de changements de qualité (n_{sw}). Par exemple, dans le scénario C, la méthode Q2ABR améliore de 78% les performances de HLS, de 62% les performances de BOLA et de 57% les performances de OSCAR, malgré les mauvaises conditions de réseau dans ce cas.

D'après la figure (5.12) on observe que la solution Q2ABR offre le meilleur taux de transfert de données par rapport aux autres algorithmes testés (HLS, BOLA et OSCAR) dans chaque scénario testé. Lorsque les conditions du réseau sont mauvaises, comme dans le scénario C, la solution Q2ABR réduit le coût d'utilisation du réseau de presque la moitié par rapport aux autres algorithmes testés (HLS [66%], BOLA [49%] et OSCAR [65%]). La solution Q2ABR réduit le trafic transféré de 20% à 35% dans le scénario B et de 37% à 42% dans le scénario A. Ces résultats sont justifiés par le choix de la représentation vidéo initiale (qualité du premier segment), qui est basée sur la première mesure de la capacité du réseau. En fait, la solution Q2ABR tente d'utiliser la même

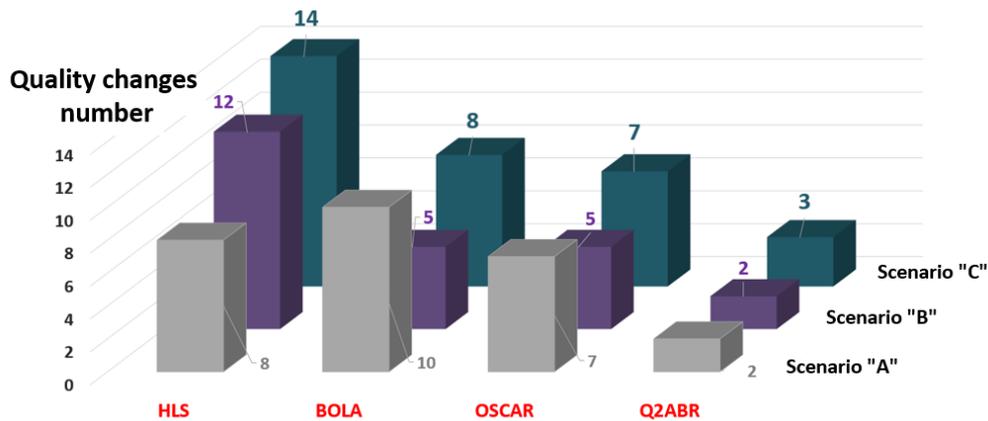


FIGURE 5.11 – Résultats de la comparaison des méthodes ABR en termes de nombre de changements de qualité n_{sw} .

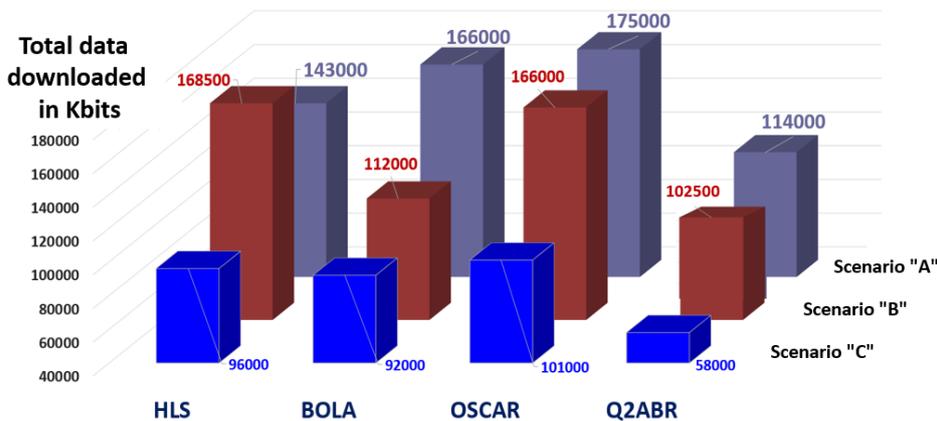


FIGURE 5.12 – Résultats de la comparaison des méthodes de RAP en termes de données totales téléchargées (Kbits)

qualité à condition que l'utilisateur soit satisfait et que le tampon ait un taux de remplissage adéquat. Cela permet de réduire le coût d'utilisation du réseau, ce qui conduit à optimiser l'utilisation des installations des opérateurs.

Enfin, la mesure des scores MOS moyens et de l'écart-type est nécessaire pour déterminer la satisfaction globale des utilisateurs de vidéos, étant donné leur nature subjective. La satisfaction des utilisateurs (QoE) sera meilleure avec des scores MOS moyens élevés et une faible variation (écart-type). La figure (5.13) illustre les scores MOS moyens calculés et la variation résultant des différentes méthodes ABR à l'aide des trois scénarios testés.

On observe que la solution Q2ABR fournit de moins bons résultats que les autres solutions (HLS, BOLA et OSCAR) dans les scénarios testés. Ceci est justifié par le fait que les algorithmes testés (HLS, BOLA, et OSCAR) favorisent la sélection de représentations vidéo élevées qui maximisent les scores MOS moyens. Cependant, nous notons que le MOS moyen prédit par notre solution Q2ABR fournit un niveau acceptable de QoE avec un MOS moyen (n_{mos}) supérieur à 3,5 dans le scénario C (mauvaises conditions de réseau) et 4 (bonne qualité selon ITU [July 2016. p. 1-42]) dans les scénarios A et B (bonnes conditions de réseau). Ce résultat nous permet de garantir notre objectif de fournir une qualité d'expérience utilisateur adéquate et de ne pas maximiser cette qualité toujours dans une vision pragmatique.

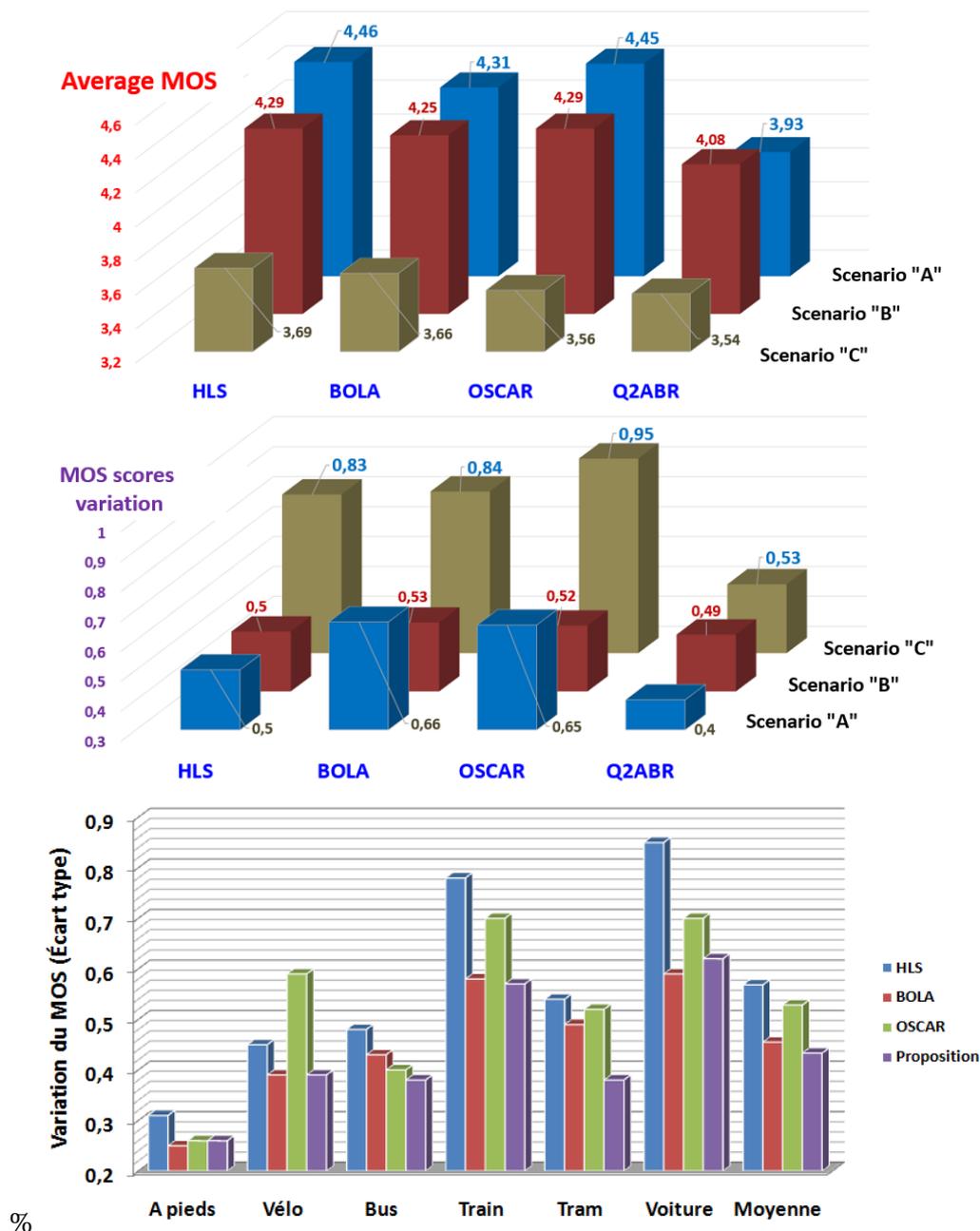


FIGURE 5.13 – Résultats de la comparaison des méthodes RAB en termes de moyenne et de variation des scores MOS prédits

5.8 Synthèses

5.8.1 Difficultés rencontrées

Encore une fois, le manque de données de qualité s'est avéré être un véritable frein au développement de ce type de travaux. Une discussion autour de la reproductibilité de la recherche scientifique devrait permettre de lever cette difficulté. La seconde difficulté fut de choisir les exemples de contrôle adaptatif à mettre en oeuvre afin de valider nos approches. En effet, il est rare que les verrous auxquels sont confrontés les opérateurs soit publiés. Il en résulte un temps incompréhensible de recherche et de discussion afin d'opter pour les exemples pertinents

5.8.2 Forces et faiblesses

L'intérêt principal de nos modèles réside dans leurs simplicité. En effet, les approches mises en oeuvre, basées sur des méthodes d'apprentissage par renforcement, nécessitent très peu de données. De plus, l'architecture modulaire employée permet de faire évoluer le système en fonction des usages.

Néanmoins, force est de constater qu'à ce jour il est très difficile de valider en production ce type d'approche ou à minima, avoir un retour constructif de la part des acteurs du terrain en plus de l'évaluation par les pairs. Il serait peut être intéressant de penser à un nouvel espace d'échange, ouvert, où les productions scientifiques publiées seraient discutées librement et publiquement par la communauté.

5.8.3 Possible extension

La première contribution présentée dans ce chapitre concerne la mise en oeuvre d'un mécanisme de routage de segment adaptatif. Ce dernier est pensé pour éviter les anomalies qui peuvent apparaître dans un réseau (par exemple, les liaisons encombrées, etc.) et répondre aux différentes exigences des usages en terme d'accords de niveau de service. Néanmoins, il est illusoire de penser qu'une simple action au niveau de routage permettrait de rendre le système totalement autonome et résilient. En effet, au delà d'une approche réactive, il est nécessaire de détecter les causes profondes du problème et y remédier définitivement. Cela implique de : i) réfléchir sur un mécanisme d'analyse des anomalies plus élaborer, ii) proposer une multitudes de mécanismes adaptatifs capables d'agir sur tous les composants du systèmes.

Ainsi, après nous être intéressés à la couche réseau, nous avons proposé un mécanisme de contrôle de la couche applicatif sous la forme d'une nouvelle méthode ABR basée sur l'estimation subjective de l'expérience de l'utilisateur (QoE) pour garantir la meilleure qualité vidéo (débit binaire) du service de streaming vidéo. L'idée principale est de garantir une qualité d'expérience acceptable pour l'utilisateur et d'optimiser l'utilisation de la bande passante. Cependant, cette approche pourrait être étendue, en intégrant plus efficacement les contraintes de la mémoire tampon, en améliorant les mécanismes de prédictions évoluer et intégrant d'autres usages tels que les nouveaux services de vidéo immersifs, de réalité virtuelle (VR), de réalité mixte (MR) ou encore de réalité augmentée (AR).

5.8.4 Impact potentiel des travaux

Comme pour la plus part de nos contributions, nous avons pris le parti de partager nos codes sources avec la communauté. Force est de constater que ce partage a généré des échanges qui n'aurait pas pu se faire sans cette initiative.

5.8.5 Retour d'expérience

Le partage des implémentations et des codes sources n'est pas dénué d'intérêt. En effet, cela facilite la reproductibilité. Néanmoins, même si certains continuent dans la même démarche de publication libre des codes sources issues de nos travaux et respectent par conséquent la licence que nous avons mise en place, cette démarche, à mon grand regret, n'est pas encore systématique.

Il est également nécessaire d'insister sur l'intérêt et l'importance de mener ce type de travaux de recherche, à dimension plurielle, dans le cadre de projets collaboratifs et avec un ancrage industriel.

Bibliographie

February 2021, «Mininet», URL <http://mininet.org/>. 81

- February 2021, «Open network operating system (onos)», URL <https://wiki.onosproject.org/display/ONOS/ONOS>. 81
- July 2016. p. 1-42, «International telecommunication union - telecommunication (itu-t) recommendation p800.1, mean opinion score (mos) terminology», . 84, 89
- Azumi, M., T. Kurosaka et M. Bandai. 2015, «A qoe-aware quality-level switching algorithm for adaptive video streaming», dans *2015 IEEE Global Communications Conference (GLOBECOM'15)*, p. 1–5, doi :10.1109/GLOCOM.2015.7417622. 84
- Bahnasse, A., F. E. Louhab, H. A. Oulahyane, M. Talea et A. Bakali. 2018, «Novel sdn architecture for smart mpls traffic engineering-diffserv aware management», *Future Generation Computer Systems*, vol. 87, p. 115–126. 78
- Barakabitze, A. A., I.-H. Mkwawa, A. Hines, L. Sun et E. Ifeakor. 2020, «Qoemultisdn : Management of multimedia services using mptcp/sr in softwarized and virtualized networks», *IEEE Access*. 78
- C. Liu, I. B. et M. Gabbouj. February 2011. p. 169-174, «Rate adaptation for adaptive http streaming», dans *Proceedings of the Second Annual ACM Conference on Multimedia Systems (MM-Sys'11)*, ISBN 978-1-4503-0518-1. 84
- CISCO. February 2021, «Cisco white paper report : Encrypted traffic analytics», URL <https://www.cisco.com/c/dam/en/us/solutions/collateral/enterprise-networks/enterprise-network-security/nb-09-encrytd-traf-anlytcs-wp-cte-en.pdf>. 77
- cisco. January 2021a, «Cisco and vodafone showcase mobile transport networking advancements via segment routing at mobile world congress», . 77
- cisco. January 2021b, «Cisco to optimize ntt docomo's 5g mobile backhaul for simpler, more flexible and scalable network operation», . 77
- CISCO Systems, I. June 2017, «Cisco white paper : Visual networking index : Forecast and methodology, 2016-2021», , p. 1–17. 83
- Clark, D. D., C. Partridge et J. C. Ramming. 2003, «A knowledge plane for the internet», dans *In SIGCOMM*, p. 3–10. 76
- Custura, A., R. Secchi et G. Fairhurst. 2018, «Exploring dscp modification pathologies in the internet», *Computer Communications*, vol. 127, p. 86–94. 77, 78
- Dong, C., C. Zhang, Z. Lu, B. Liu et B. Jiang. 2020, «Cetanalytics : Comprehensive effective traffic information analytics for encrypted traffic classification», *Computer Networks*, vol. 176, p. 107258. 77
- Feamster, N., J. Rexford et E. Zegura. 2014, «The road to sdn : An intellectual history of programmable networks», *SIGCOMM Comput. Commun. Rev.*, vol. 44, n° 2, doi :10.1145/2602204.2602219, p. 87–98, ISSN 0146-4833. URL <https://doi.org/10.1145/2602204.2602219>. 77
- Grigoriou, E., A. A. Barakabitze, L. Atzori, L. Sun et V. Pilloni. 2017, «An sdn-approach for qoe management of multimedia services using resource allocation», dans *2017 IEEE International Conference on Communications (ICC)*, IEEE, p. 1–7. 82
- Han, B., F. Qian, L. Ji et V. Gopalakrishnan. 2016, «Mp-dash : Adaptive video streaming over preference-aware multipath», dans *Proceedings of the 12th International on Conference on Emerging Networking EXperiments and Technologies*, (CoNEXT'16), ACM, ISBN 978-1-4503-4292-6, p. 129–143. 86, 87

- He, Y.-L., X.-L. Zhang, W. Ao et J. Z. Huang. 2018, «Determining the optimal temperature parameter for softmax function in reinforcement learning», *Applied Soft Computing*, vol. 70, p. 80–85. 81
- J. Jiang, V. S. et H. Zhang. January 2014, «Improving fairness, efficiency, and stability in http-based adaptive video streaming with festive», dans *Proceedings of the 8th International Conference on Emerging Networking Experiments and Technologies*, (CoNEXT'12), ISBN 978-1-4503-1775-7, p. 97–108. 84
- Jin, C., Z. Yang, Z. Wang et M. I. Jordan. 2020, «Provably efficient reinforcement learning with linear function approximation», dans *Conference on Learning Theory*, PMLR, p. 2137–2143. 80
- Joseph, V. et G. de Veciana. 2014, «Nova : Qoe-driven optimization of dash-based video delivery in networks», dans *IEEE INFOCOM 2014 - IEEE Conference on Computer Communications*, p. 82–90. 84
- Juluri, P., V. Tamarapalli et D. Medhi. 2015, «Sara : Segment aware rate adaptation algorithm for dynamic adaptive streaming over http», dans *2015 IEEE International Conference on Communication Workshop (ICCW)*, p. 1765–1770. 84
- Koenig, S. et R. G. Simmons. 1993, «Complexity analysis of real-time reinforcement learning», dans *AAAI*, p. 99–107. 80
- Kukreja, N., R. Alvizu, A. Kos, G. Maier, R. Morro, A. Capello et C. Cavazzoni. 2016, «Demonstration of sdn-based orchestration for multi-domain segment routing networks», dans *2016 18th International Conference on Transparent Optical Networks (ICTON)*, IEEE, p. 1–4. 77, 78
- Kushwaha, A., S. Sharma, N. Bazard, A. Gumaste et B. Mukherjee. 2020, «Design, analysis, and a terabit implementation of a source-routing-based sdn data plane», *IEEE Systems Journal*. 77
- Li, Z., S. Peng, D. Voyer, C. Xie, P. Liu, Z. Qin, K. Ebisawa, S. Previdi et J. Guichard. 2020, «Problem statement and use cases of application-aware networking (apn)», Internet-Draft draft-li-apn-problem-statement-usecases-01, Internet Engineering Task Force. URL <https://datatracker.ietf.org/doc/html/draft-li-apn-problem-statement-usecases-01>, work in Progress. 77, 78
- Lin, S.-C., I. F. Akyildiz, P. Wang et M. Luo. 2016, «Qos-aware adaptive routing in multi-layer hierarchical software defined networks : A reinforcement learning approach», dans *2016 IEEE International Conference on Services Computing (SCC)*, IEEE, p. 25–33. 82
- Liotou, E., D. Tsolkas et N. Passas. 2016, «A roadmap on QoE metrics and models», dans *2016 23rd International Conference on Telecommunications (ICT'16)*, p. 1–5. 84
- M. Claeys, S. L., J. Famaey, T. Wu, L. Van et F. D. Turck. May 2013, «Design of a q-learning-based client quality selection algorithm for http adaptive video streaming», dans *Adaptive and Learning Agents Workshop*, p. 30–37. 84
- Mammeri, Z. 2019, «Reinforcement learning based routing in networks : Review and classification of approaches», *IEEE Access*, vol. 7, p. 55 916–55 950. 80
- McLean, P. 2009, «Apple launches http live streaming standard in iphone 3.0», *AppleInsider*. Retrieved 15 October 2011. 84, 88
- Menkovski, V. et A. Liotta. 2013, «Intelligent control for adaptive video streaming», dans *2013 IEEE International Conference on Consumer Electronics (ICCE'13)*, p. 127–128. 84, 85

- Mushtaq, M., B. Augustin et A. Mellouk. 2015, «Regulating QoE for adaptive video streaming using BBF method», dans *2015 IEEE International Conference on Communications (ICC'15)*, p. 6855–6860. [84](#)
- Peng, S., J. Mao, R. Hu et Z. Li. 2020, «Demo abstract : Apn6 : Application-aware ipv6 networking», dans *IEEE INFOCOM 2020-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPs)*, IEEE, p. 1330–1331. [77](#), [78](#)
- Priyadarshana, W. J. et G. Sofronov. 2015, «Multiple break-points detection in array cgh data via the cross-entropy method», *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 12, n° 2, p. 487–498, ISSN 1545-5963. [85](#), [86](#)
- R. Bertrand, P. M. Fhima et A. Guillin. 2010. p 1-26, «Off-line detection of multiple change points with the filtered derivative with p-value method», *Sequential Analysis Journal*. [84](#)
- Rego, A., S. Sendra, J. M. Jimenez et J. Lloret. 2019, «Dynamic metric ospf-based routing protocol for software defined networks», *Cluster Computing*, vol. 22, n° 3, p. 705–720. [78](#)
- RY. Rubinstein. May 1997, «Optimization of Computer Simulation Models with Rare Events», *European Journal of Operational Research*, vol. Volume 99, n° 1, p. p. 89–112, ISSN 0377-2217. [85](#)
- Spiteri, K., R. Urgaonkar et R. K. Sitaraman. 2016, «Bola : Near-optimal bitrate adaptation for online videos», dans *The 35th Annual IEEE International Conference on Computer Communications (INFOCOM'16)*, p. 1–9. [84](#), [88](#)
- Sterritt, R. 2004, «Autonomic networks : engineering the self-healing property», *Engineering Applications of Artificial Intelligence*, vol. 17, n° 7, doi :<https://doi.org/10.1016/j.engappai.2004.08.028>, p. 727–739, ISSN 0952-1976. URL <https://www.sciencedirect.com/science/article/pii/S0952197604001113>, autonomic Computing Systems. [76](#)
- Sun, Y., X. Yin, J. Jiang, V. Sekar, F. Lin, N. Wang, T. Liu et B. Sinopoli. 2016, «Cs2p : Improving video bitrate selection and adaptation with data-driven throughput prediction», dans *Proceedings of the 2016 ACM SIGCOMM Conference (SIGCOMM'16)*, ACM, p. 272–285. [84](#)
- Temlyakov, V. 2015, «Greedy approximation in convex optimization», *Constructive Approximation*, vol. 41, n° 2, p. 269–296. [81](#)
- Tong, V. February 2021, «Service-centric segment routing using reinforcement learning», URL <https://github.com/vanvantong/rl-sr>. [82](#)
- Tong, V., H. A. Tran, S. Souihi et A. Mellouk. 2020, «Service-centric segment routing mechanism using reinforcement learning for encrypted traffic», dans *2020 16th International Conference on Network and Service Management (CNSM)*, IEEE. [78](#)
- Tran, H. A., S. Hoceini, A. Mellouk, J. Perez et S. Zeadally. 2013, «Qoe-based server selection for content distribution networks», *IEEE Transactions on Computers*, vol. 63, n° 11, p. 2803–2815. [81](#)
- T.Y Huang, N. M. M. T., R. Johari et M. Watson. August 2014, «A buffer-based approach to rate adaptation : Evidence from a large video streaming service», dans *Proceedings of the 2014 ACM Conference on SIGCOMM (SIGCOMM'14)*, ISBN 978-1-4503-2836-4, p. 187–198. [84](#), [88](#)
- V. Krishnamoorthi, E. H., N. Carlsson et E. Petajan. June 2017, «Buffest : Predicting buffer conditions and real-time requirements of http(s) adaptive streaming clients», dans *ACM SIGMM Conference on Multimedia Systems (MMSys'17)*, p. 1–12. [84](#)

Yap, K.-K., M. Motiwala, J. Rahe, S. Padgett, M. Holliman, G. Baldus, M. Hines, T. Kim, A. Narayanan, A. Jain et collab.. 2017, «Taking the edge off with espresso : Scale, reliability and programmability for global internet peering», dans *Proceedings of the Conference of the ACM Special Interest Group on Data Communication*, p. 432–445. [78](#)

Zahran, A. H., J. Quinlan, D. Raca, C. J. Sreenan, E. Halepovic, R. K. Sinha, R. Jana et V. Gopalakrishnan. May 2016, «Oscar : An optimized stall-cautious adaptive bitrate streaming algorithm for mobile networks», dans *Proceedings of the 8th International Workshop on Mobile Video (MoVid'16)*, p. 1–6. [84](#), [86](#), [88](#)

Zhou, C., C. W. Lin, X. Zhang et Z. Guo. 2013, «Buffer-based smooth rate adaptation for dynamic http streaming», dans *2013 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference*, p. 1–9. [84](#), [88](#)

Chapitre 6

Conclusion Générale

*« The closer you look at something,
the more complex it seems to be. »*

Vinton Cerf

Sommaire

6.1 Contributions et retours d'expérience	98
6.2 Autres travaux	98
6.3 Méthode de travail : organisation agile pour la recherche	100
6.3.1 Scrum'RLite : Une Organisation Agile	100
6.3.2 Une recherche libre et reproductible	101
6.4 Défis scientifiques et perspectives	102

6.1 Contributions et retours d'expérience

La croissance continue du trafic de données, l'émergence de la virtualisation des réseaux ainsi que l'utilisation sans cesse croissante d'équipements mobiles dans l'environnement réseau moderne ont mis en lumière les nombreux problèmes inhérents à l'architecture conventionnelle de l'Internet. Ainsi, la tâche de gestion et de contrôle des informations provenant d'un nombre croissant d'appareils connectés devenait de plus en plus complexe et spécialisée.

En effet, l'infrastructure réseau traditionnelle est considérée comme très rigide et statique. Elle a été conçue à l'origine pour un type de trafic particulier, à savoir des contenus monotones en texte, ce qui la rend peu adaptée aux flux multimédia interactifs et dynamiques d'aujourd'hui générés par des utilisateurs de plus en plus exigeants et mobiles. Parallèlement aux besoins liés au multimédia, l'émergence récente de l'Internet des Objets (IoT) a permis la création de nouveaux services avancés avec des exigences de communication plus strictes afin de prendre en charge des cas d'utilisation innovants. En particulier, la santé connectée est un cas typique d'utilisation de l'IoT où les services de soins de santé fournis à des patients distants (e.g. diagnostic, chirurgie, dossiers médicaux) sont particulièrement stricts au regard du délai, de la qualité et de la confidentialité. Ces données sensibles et ce trafic critique ne sont guère pris en charge par les réseaux traditionnels.

De plus, dans l'architecture traditionnelle où la logique de contrôle est purement distribuée et localisée, la résolution d'un problème de réseau spécifique ou le réglage d'une stratégie de réseau particulière nécessite d'agir séparément sur les périphériques concernés et de modifier manuellement leur configuration. Dans ce contexte, la croissance actuelle du nombre d'appareils et de données a accru les problèmes d'évolutivité en rendant ces interventions humaines et ces opérations de réseau plus dures et plus sujettes aux erreurs.

Cela dit, cet énorme fossé entre, d'une part, les progrès réalisés dans les technologies informatiques et logicielles et, d'autre part, l'infrastructure de réseau sous-jacente traditionnelle, non évolutive et difficile à gérer. [Bannour et collab. \[2017a,b\]](#) ont souligné le besoin d'une plate-forme de réseau automatisée et adaptable en continu qui facilite les opérations du réseau et réponde aux besoins de l'IoT. Dans ce contexte, plusieurs stratégies de recherche ont été proposées pour intégrer des approches automatiques et adaptatives dans l'infrastructure actuelle afin de relever les défis de l'évolutivité, de la fiabilité et de la disponibilité du trafic en temps réel, garantissant ainsi la QoE de l'utilisateur.

6.2 Autres travaux

Dans cette section nous présentons certains travaux effectués durant les 9 dernières années, en lien avec ceux présentés dans ce manuscrit, mais que nous avons délibérément choisi de ne pas intégrer dans le présent document dans la mesure où il ne s'agissait pas de travaux faits en collaboration avec des doctorants ou qui relevaient plus d'un enjeu technologique et d'ingénierie que d'un verrou scientifique.

La première de ces contributions concerne l'interopérabilité. Il s'agit de la capacité des contrôleurs SDN à échanger des informations et à coordonner un fonctionnement sans faille. Dans une architecture SDN générale, en plus de l'interface nord et de l'interface sud qui servent à échanger des informations avec le plan d'application et le plan de données, l'interface est-ouest peut être mise en oeuvre pour la communication entre les contrôleurs distribués. Cependant, contrairement à l'interface sud qui est standardisée OpenFlow, l'interface est-ouest n'a pas reçu l'attention nécessaire de la part de la communauté des chercheurs pour offrir le même cadre. Il est indéniable que les efforts de certains groupes de recherche ont donné quelques résultats, voir [\[Brockelsby et Dutta, 2021; Gerola et collab., 2016\]](#). Cependant, ils ont juste résolu le problème d'interopérabilité pour un réseau SDN homogène avec le même type de contrôleur fourni par un fournisseur. Récemment, deux interfaces est-ouest [\[Almadani et collab., 2021; Yu et collab., 2020\]](#) ont été proposées pour les réseaux SDN hétérogènes avec différents types de contrôleurs. Néanmoins, ils ont ap-

pliqué le mécanisme de réplication active en diffusant chaque nouvelle information de mise à jour à tous les autres contrôleurs. En conséquence, la méthode de diffusion dégrade les performances du réseau. Nous avons proposé une interface est-ouest appelée SINA, qui aide les contrôleurs SDN hétérogènes à mettre à jour les états du réseau. L'interface SINA (SDN Inter-Domains Network Application) proposée est une interface est-ouest permettant le partage de l'état du réseau entre des domaines SDN hétérogènes et dont l'architecture est entièrement personnalisable. Concrètement, la politique de notification peut être adaptée en fonction de la nature de l'information, du niveau de criticité ou des conditions du réseau. En outre, pour sécuriser la transmission des informations, une autorisation de transfert de données basée sur le protocole OAuth2 (open authorization) (<https://oauth.net/2/> consulté en 2020) peut être activée. Ce dernier peut prendre en charge différents jetons pour fournir plusieurs niveaux d'autorisation et les adapter dynamiquement en fonction des besoins. Afin de prouver l'ouverture de SINA, nous le déployons dans trois contrôleurs SDN différents et largement utilisés : ONOS [Berde et collab., 2014], OpenDaylight [Medved et collab., 2014], et Faucet [Bailey et Stuart, 2016]. La figure 6.1 illustre l'implémentation de SINA dans ces trois contrôleurs ainsi que les différentes interactions.

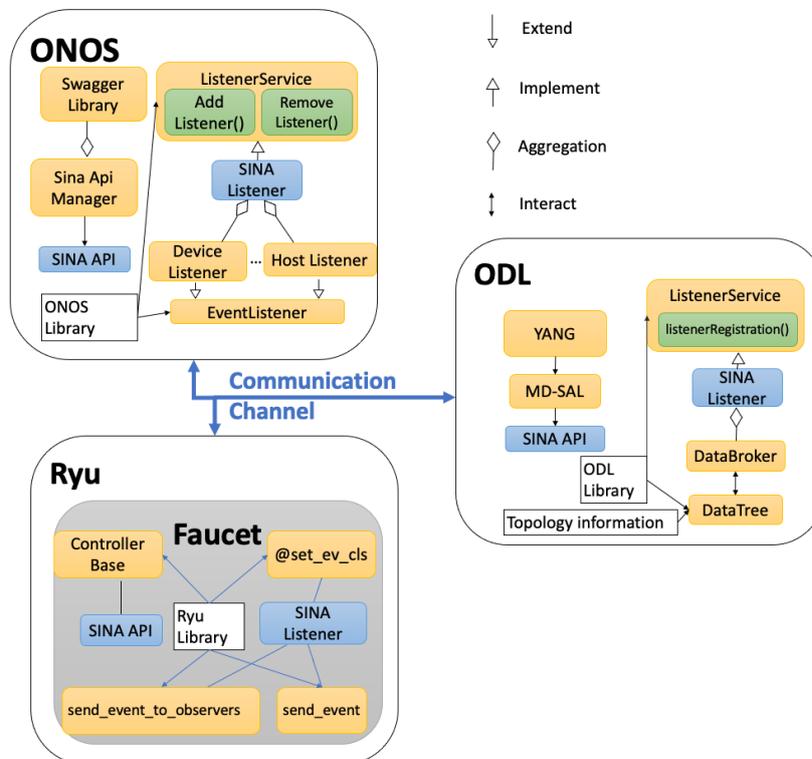


FIGURE 6.1 – Implémentation de SINA

La seconde contribution concerne la sécurité des échanges dans entre contrôleurs SDN. En effet, malgré toute l'attention que lui apporte la communauté scientifique, le SDN est encore un nouveau paradigme, il est donc difficile d'éviter tout défaut en termes d'évolutivité, de fiabilité et de sécurité. C'est d'autant plus vrai pour une architecture déployée et physiquement distribuée et logiquement centralisée. Cette architecture, basée sur la collaboration de plusieurs contrôleurs SDN, introduit de nouveaux problèmes de sécurité. En effet, un contrôleur malveillant ou défaillant peut provoquer de graves perturbations dans le réseau en alimentant ses pairs avec de mauvaises connaissances. Le développement de la Blockchain permet de résoudre ces problèmes. La Blockchain est une technologie émergente qui utilise des mécanismes de consensus pour effectuer la validation des données sans s'appuyer sur un tiers de confiance. Une fois que les données ont été ajoutées à un bloc, elles sont immuables, de sorte que personne ne peut les modifier, ce qui garantit la fiabilité des informations. Tous les éléments d'information de la Blockchain sont sto-

ckés publiquement et peuvent être consultés par n'importe qui, ce qui garantit la cohérence entre les membres du réseau et la traçabilité de toutes les transactions. La Blockchain est donc une excellente solution pour améliorer le fonctionnement du réseau SDN. De nombreux articles traitent et étudient cette question : [Aujla et collab., 2020], [Jiasi et collab., 2019]. [Nam Nguyen et collab., 2021]. Dans ce contexte, nous avons travaillé sur un modèle utilisant la Blockchain pour améliorer le fonctionnement du réseau SDN. En outre, nous avons proposé également quelques solutions pour améliorer la Blockchain, en termes d'évolutivité qui concernent notamment le modèle de consensus.

6.3 Méthode de travail : organisation agile pour la recherche

Dans cette section, je vais brièvement exposer mes conclusions concernant une méthode de travail qui pourrait contribuer à améliorer le fonctionnement d'un groupe de recherche. Je ne prétends pas que ces méthodes soient généralisables dans l'état. Il s'agit uniquement d'un retour basé sur ma propre expérience.

6.3.1 Scrum'RLite : Une Organisation Agile

Les chercheurs et les enseignants chercheurs sont confrontés à un certains nombres de challenges dans leurs pratiques quotidiennes. En effet, nous devons jongler entres différentes tâches dans la semaine et parfois dans la même journée, nous devons aussi composer avec une chaîne de décision qui allonge les délais, un manque d'information ou des informations dispersées et non organisées. En fait, en plus de notre coeur de métier, il nous est demandé de plus en plus d'être de véritables Chef-de-projets, métier pour lequel nous ne sommes certainement pas formés. Il en résulte une importante perte de temps, un manque d'efficacité et une certaine frustration voire une démotivation. Il m'a donc paru important de partager mes réflexions à ce sujet.

Si nous faisons le parallèle avec le monde de l'entreprise, il peut être intéressant d'adapter certaines de leurs méthodes au monde académique. Les méthodes de gestions de projet Agile et particulièrement la méthode SCRUM [Schwaber et Sutherland, 2017] sont particulièrement adaptées. Dans ce contexte, un travail de réflexion important a été mené par May et Runyon [2019]. Cette réflexion a donnée naissance à une approche, intitulée LabScrum, qui est une tentative d'adaptation de la méthode SCRUM dans le contexte académique. Bien qu'intéressante, cette approche reste à mon avis, trop complexe et inadaptée aux spécificité des équipes de recherche en France; ce qui m'a amené à réfléchir à une approche simplifiée que j'ai appelé Scrum'RLite. Dans cette section je ne vais pas développer les détails de la méthode mais uniquement les éléments de différence avec LabScrum.

Trop de SCRUM, tue le SCRUM

Pour illustrer cette idée, quoi de plus simple que de paraphraser Montesquieu : *"le mieux est l'ennemi mortel du bien"*. En effet, la meilleure manière de brider l'efficacité d'une méthode de gestion de projets est d'établir trop de règles. Par exemple, définir chaque thèse de doctorat comme un projet SCRUM séparé me semble constituer un non-sens. L'objectif escompté est justement de réduire l'effet silos en permettant et en facilitant les collaborations. Ainsi cette méthode de travail doit permettre une totale transparence et un échange permanent entre les individus et les projets.

Durée d'un sprint

Un sprint est une itération dans la méthode SCRUM. Sa durée peut varier en fonction du contexte (par exemple, la préconisation de LabScrum est de 2 semaines dans la majorité des cas). Il est évidemment impossible de répondre à la question "quelle est la durée idéale d'un sprint?",

mais une durée d'un mois (en choisissant par exemple, le premier lundi du mois comme journée de revue de sprint) m'est parue adaptée d'après mon expérience. Cela laisse suffisamment de temps aux chercheurs pour la réflexion.

Les acteurs et les rôles

La méthode SCRUM définit principalement 3 rôles : le "Product Owner (PO)", le "Scrum Master" et le "Development Team Member".

- Le Product Owner (PO), est le lien entre la partie métier et la partie technique d'un projet. Bien qu'intéressant notamment dans les projets collaboratifs, ce rôle n'est pas conservé dans la méthode Scrum'RLite afin de réduire la complexité de la méthode.
- Le "Development Team Member" est la cheville ouvrière de la méthode SCRUM. Il est censé s'auto-organiser afin d'accomplir les tâches qu'il s'est fixé. En l'occurrence, chaque membre de l'équipe de recherche, quelque soit son rôle, fait partie de cette catégorie.
- Le "Scrum Master", est chargé d'assurer que la méthode Scrum est comprise et correctement mis en œuvre. Même si certains sont contre l'idée d'un Scrum Master tournant (pour des raisons légitimes), la méthode Scrum'RLite adopte justement cette idée. Pourquoi? Afin, encore une fois, de favoriser la collaboration. En effet cela oblige les membres à prendre connaissance des travaux envisagés ou effectués dans l'équipe. Néanmoins, ce rôle sera réservé aux membres permanents ainsi qu'aux doctorants en fin de thèse afin de les préparer progressivement à leurs futures carrières.

Daily Scrum

La Daily Scrum est une réunion quotidienne et courte (ne dépassant pas les 15 minutes) qui permet à chacun d'exposer son planning mais c'est surtout l'occasion de lever des alertes. Cela permet d'éviter l'effet goulot d'étranglement (un doctorant par exemple qui aurait besoin d'une signature de son directeur de thèse, ou un membre de l'équipe qui se trouverait face à un bug bloquant dans une plateforme d'expérimentation, ...). Bien que cette réunion soit généralement programmée le matin, dans le contexte académique il est plus simple dans cette situation de la programmer en début d'après-midi et avant la reprise.

6.3.2 Une recherche libre et reproductible

Dans cette section, je vais brièvement exposer mes réflexions sur la reproductibilité de la recherche. Cette dernière est un principe fondamental de la science, mais force est de constater que certaines études scientifiques ne peuvent être reproduites. Cela est dû à plusieurs facteurs : absence des données d'apprentissages, absence de codes sources, difficultés à configurer correctement les systèmes, difficultés à reproduire les conditions, problèmes méthodologiques, La question se pose donc sur la manière pour y remédier sachant que la reproductibilité devrait être considérée comme un pré-requis pour toute étude scientifique. Je crois pour ma part qu'une méthode incitative permettrait d'améliorer cet aspect. Néanmoins, dans cette section, je vais m'intéresser uniquement à une méthodologie à mettre en place pour faciliter la reproductibilité.

Des données ou, à défaut, des modèles pré-entraînés

La reproductibilité commence dès la phase de planification, avec de bonnes pratiques de gestion des données. En effet, il est souvent très difficile de reproduire une recherche lorsque les données sont désorganisées ou lorsqu'il manque des méta-informations sur la manière dont ces données ont été obtenues. Il est donc important de : i) prioriser les formats standard de données, ii) prêter une attention particulière au noms des colonnes et des variables, iii) proposer systématiquement des méta-données afin de faciliter la ré-utilisation.

Dans le cas où le partage des données est rendu impossible par les contraintes techniques ou réglementaires, il reste toujours la possibilité de partager un modèle pré-entraîné. Ce dernier, à défaut de proposer une solution de reproductibilité, offre à la communauté la possibilité de comparer les modèles voire de valider une nouvelle fois les modèles sur de nouvelles données collectées

Des données oui mais versionnées

Tout au long du processus d'acquisition des données, un processus de contrôle de version devrait être utilisé afin de conserver des changements survenus au cours de la vie d'un projet. Les systèmes de contrôle de version, par exemple ceux basés sur Git, enregistrent les modifications apportées à un fichier ou à un ensemble de fichiers au fil du temps. Ces systèmes sont relativement faciles à configurer et à utiliser et permettent de récupérer les données utilisées à un moment précis (par exemple, lors de la soumission d'un papier), même si celles-ci sont mises à jour ultérieurement.

6.4 Défis scientifiques et perspectives

Ce manuscrit a été l'occasion de présenter une approche pragmatique pour mettre en oeuvre un réseau intelligent. Dans un premier temps, j'ai présenté les idées que nous avons développées pour décentraliser le plan de contrôle SDN dans les réseaux à large échelle. Ces idées, bien que prometteuses, peuvent être étendues.

- Tout d'abord, les stratégies de placement des contrôleurs proposées pourraient être encore améliorées en incluant *une politique dynamique de placement des contrôleurs*. Cette dernière devrait prendre en compte la nature dynamique du réseau, comme la charge du réseau ou une topologie changeante.
- Le modèle de cohérence adaptative proposé réduit la surcharge inter-contrôleurs et ajuste le niveau de cohérence au moment de l'exécution, afin d'obtenir un compromis. Il est toutefois nécessaire de considérer le caractère variable et très dynamique que peut avoir un réseau. Ce dernier peut en effet augmenter d'une façon significative la complexité du problème. Une autre approche possible consiste à attribuer différents niveaux de cohérence aux différents contrôleurs (cohérence granulaire) en fonction des exigences de l'application. De la même manière. Il serait également intéressant d'adapter le degré de réplication pour chaque contrôleur afin de diffuser les connaissances dans des zones géographiques spécifiques en fonction des besoins.
- Ensuite, il convient de considérer le problème du placement des contrôleurs et le problème du partage des connaissances entre les contrôleurs SDN distribués de manière corrélée. Certes, cela représentera un défi supplémentaire mais non moins nécessaire. Par exemple, la minimisation des latences entre contrôleurs dans le processus de placement des contrôleurs réduit le coût des communications entre contrôleurs et améliore la cohérence et les performances du réseau.
- Enfin, il peut être intéressant de considérer l'aspect sécurité dans la communication inter-contrôleurs afin de protéger le système contre les menaces potentielles auxquelles est confronté le plan de contrôle SDN. Un exemple simple de ces menaces est une réplique de contrôleur SDN malveillante qui envoie des données erronées pour compromettre le système en nuisant à un service ou un réseau particulier, ou en favorisant ses actions au détriment du reste des répliques de contrôleur. Une solution potentielle consiste à utiliser une blockchain pour fournir une garantie de non-répudiation et de non-altération (intégrité) des données. Cette blockchain peut également être utilisée pour stocker des contrats intelligents. Ces contrats sont des programmes qui contrôlent l'autorisation des échanges de données entre les parties sous certaines conditions. En particulier, les contrats intelligents peuvent permettre un

contrôle d'accès plus fin et plus précis des connaissances à partager entre les répliques de contrôleur SDN en établissant des règles élaborées (par exemple en autorisant l'accès à une connaissance particulière uniquement pour des répliques de contrôleur spécifiques).

Dans un second temps on s'est intéressé aux connaissances nécessaires aux opérateurs afin de contrôler et gérer les infrastructures réseaux. Ce point en particulier demeure encore très ouvert dans la mesure où les connaissances pouvant être induites ou acquises dans un réseau ne peuvent qu'augmenter.

- La première perspective consiste à poursuivre le travail sur l'estimation de la qualité de l'expérience. En premier lieu, il est important de travailler sur la collecte de données, étape essentielle pour améliorer la précision des modèles d'estimation de QoE. L'un des problèmes réside dans le peu de données d'apprentissage, et particulièrement pour les exemples pouvant aboutir à des situations contradictoires (cas d'un grand écart entre les valeurs de MOS données par les testeurs pour une combinaison de facteurs donnée). Au lieu de se contenter de combinaisons aléatoires dans les tests, il serait intéressant de réfléchir à une méthode permettant de favoriser l'utilisation des combinaisons de facteurs qui mènent à des situations contradictoires. Autrement dit, utiliser les valeurs de facteurs qui détériorent les performances de l'estimation de la QoE. L'une des pistes prometteuses à explorer consiste à utiliser l'apprentissage actif (*Active Learning*) pour la réalisation de campagnes de test.
- Ensuite, il est nécessaire de continuer les efforts liés à l'identification des applications dans le cas d'un trafic chiffré. Ces approches utilisent des algorithmes ML supervisés pour classer le trafic réseau et nécessitent de grands ensembles de données étiquetées dont la collecte est coûteuse et chronophage. Par conséquent, il serait intéressant de considérer des approches de ML semi-supervisées. L'objectif est de pré-entraîner un modèle sur un grand ensemble de données non étiquetées et de transférer les poids appris à un nouveau modèle qui est réentraîné avec un petit ensemble de données étiquetées (apprentissage par transfert).

Enfin, ce manuscrit a présenté nos efforts pour la mise en oeuvre d'un contrôle adaptatif du réseau. À cet effet, nous nous sommes essentiellement concentrés sur 2 services : le routage et la diffusion vidéo adaptative. Il serait intéressant de considérer d'autres mécanismes pouvant agir sur la QoE usager. Par exemple, nous pourrions agir sur le mécanisme de gestion de la congestion dans la couche transport en proposant un algorithme BBR (Bottleneck Bandwidth and RTT) adaptatif pour QUIC afin de s'adapter aux changements inattendus des environnements réseau. Plus largement, il est important de considérer les différents mécanismes de contrôle adaptatif de manière corrélée et non plus comme des approches indépendantes. Dans ce contexte, il serait intéressant de considérer un positionnement par rapport à de nouvelles architectures comme SCION (Scalability, Control, and Isolation On Next-Generation Networks) voire un rapprochement. En effet, SCION suscite de plus en plus d'intérêt chez des opérateurs, les fournisseurs de services Internet et les entreprises et les réflexions pour sa possible mise en production se sont considérablement accélérées ces derniers temps.

Bibliographie

- Almadani, B., A. Beg et A. S. H. Mahmoud. 2021, «Dsf : A distributed sdn control plane framework for the east/west interface», *IEEE Access*, vol. 9, p. 26 735–26 754. [98](#)
- Aujla, G. S., M. Singh, A. Bose, N. Kumar, G. Han et R. Buyya. 2020, «Blocksdn : Blockchain as a service for software defined networking in smart city applications», *IEEE Network*, vol. 34, n° 2, p. 83–91. [100](#)

- Bailey, J. et S. Stuart. 2016, «Faucet : Deploying sdn in the enterprise», *Commun. ACM*, vol. 60, n° 1, doi :10.1145/3009828, p. 45–49, ISSN 0001-0782. URL <https://doi.org/10.1145/3009828>. 99
- Bannour, F., S. Souihi et A. Mellouk. 2017a, «Distributed sdn control : Survey, taxonomy, and challenges», *IEEE Communications Surveys & Tutorials*, vol. 20, n° 1, p. 333–354. 98
- Bannour, F., S. Souihi et A. Mellouk. 2017b, «Software-defined networking : A self-adaptive consistency model for distributed sdn controllers», *RESCOM 2017*. 98
- Berde, P., M. Gerola, J. Hart, Y. Higuchi, M. Kobayashi, T. Koide, B. Lantz, B. O’Connor, P. Radoslavov, W. Snow et G. Parulkar. 2014, «Onos : Towards an open, distributed sdn os», dans *Proceedings of the Third Workshop on Hot Topics in Software Defined Networking*, HotSDN ’14, Association for Computing Machinery, New York, NY, USA, ISBN 9781450329897, p. 1–6, doi :10.1145/2620728.2620744. URL <https://doi.org/10.1145/2620728.2620744>. 99
- Brockelsby, W. et R. Dutta. 2021, «Traffic analysis in support of hybrid sdn campus architectures for enhanced cybersecurity», dans *2021 24th Conference on Innovation in Clouds, Internet and Networks and Workshops (ICIN)*, p. 41–48, doi :10.1109/ICIN51074.2021.9385530. 98
- Gerola, M., F. Lucrezia, M. Santuari, E. Salvadori, P. L. Ventre, S. Salsano et M. Campanella. 2016, «Icona : A peer-to-peer approach for software defined wide area networks using onos», *2016 Fifth European Workshop on Software-Defined Networks (EWSDN)*, p. 37–42. 98
- Jiasi, W., W. Jian, L. Jia-Nan et Z. Yue. 2019, «Secure software-defined networking based on blockchain», *arXiv*. 100
- May, L. et T. Runyon. 2019, «Labscrum : A case study for agility in academic research labs», . 100
- Medved, J., R. Varga, A. Tkacik et K. Gray. 2014, «Opendaylight : Towards a model-driven sdn controller architecture», *Proceeding of IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks 2014*, p. 1–6. 99
- Nam Nguyen, H., H. Anh Tran, S. Fowler et S. Souihi. 2021, «A survey of blockchain technologies applied to software-defined networking : Research challenges and solutions», *IET Wireless Sensor Systems*, vol. 11, n° 6, p. 233–247. 100
- Schwaber, K. et J. Sutherland. 2017, «The scrum guide™», . 100
- Yu, H., H. Qi et K. Li. 2020, «Wecan : an efficient west-east control associated network for large-scale sdn systems», *MONET*, vol. 25, n° 1, doi :10.1007/s11036-018-1194-9, p. 114–124. URL <https://doi.org/10.1007/s11036-018-1194-9>. 98

Chapitre 7

Articles co-rédigés

(Doctorants, étudiants et Post-Docs supervisés en souligné)

Publications de rang A

1. Hoang, N.-T., H.-N. Nguyen, H.-A. Tran et S. Souihi. 2022, «A novel adaptive east&west interface for a heterogeneous and distributed sdn network», Electronics, vol. 11, no 7, doi:10.3390/electronics11070975, ISSN 2079-9292, Impact Factor : 2.69.
2. Nam Nguyen, H., Anh Tran, H., Fowler, S. and Souihi, S. (2021). "A survey of Blockchain technologies applied to software-defined networking : Research challenges and solutions". IET Wireless Sensor Systems, 11(6), 233–247, Impact Factor : 2.617.
3. Naloufi M, Lucas FS, Souihi S, Servais P, Janne A, Wanderley Matos De Abreu T. Evaluating the Performance of Machine Learning Approaches to Predict the Microbial Quality of Surface Waters and to Optimize the Sampling Effort. Water. 2021 ; 13(18) :2457. <https://doi.org/10.3390/w13182> Impact Factor : 3.53.
4. Tong, V., Souihi, S., Tran, H. A. and Mellouk, A. (2021). "SDN-Based Application-Aware Segment Routing for Large-Scale Network". IEEE Systems Journal, Impact Factor : 3.931.
5. F Bannour, S. Souihi, A. Mellouk : "Adaptive distributed SDN controllers : Application to Content-Centric Delivery Networks". Future Gener. Comput. Syst. 113 : 78-93 (2020), Impact Factor : 6.125.
6. S. Souihi, S. Bitam, A. Mellouk, T. Abreu, S. Hoceini, S. Fowler, S. Medileh, S. De, A. Shami : "Smart communications for autonomous systems in network technologies". Int. J. Communication Systems 33(10) (2020), Impact Factor : 1.319.
7. L. Amour, S. Souihi, A. Mellouk, M. S. Mushtaq. "Q2ABR : QoE-aware adaptive video bit rate solution". Int. J. Communication Systems 33(10) (2020), Impact Factor : 1.319.
8. H. A. Tran, S. Souihi, D. Tran, A. Mellouk. "MABRESE : A New Server Selection Method for Smart SDN-Based CDN Architecture". IEEE Communications Letters 23(6) : 1012-1015 (2019), Impact Factor : 3.419.
9. F Bannour, S. Souihi, A. Mellouk. "Distributed SDN Control : Survey, Taxonomy, and Challenges". IEEE Communications Surveys and Tutorials 20(1) : 333-354 (2018), Impact Factor : 23.7.

10. S. Souihi, S. Hoceini, A. Mellouk, B. Augustin, N. A. Saadi. "SMILAY : An information flow management framework for microgrid applications". IEEE Communications Magazine 51(1) : 120-126 (2013), Impact Factor : 9.03.

Publications de rang B

1. Balachandran, T., T. Abreu, M. Naloufi, S. Souihi, F. Lucas et A. Janne. 2022, «IoT and transfer learning based urban river quality prediction», dans 2022 IEEE Global Communications Conference : IoT and Sensor Networks (Globecom 2022 IoTSN), Rio de Janeiro, Brazil.
2. Do, Q. H., S. Souihi, V. Tong, H. A. Tran et S. Mhadhbi. 2022, «How tezos blockchain can meet IoT?», dans 2022 IEEE Global Communications Conference : Next-Generation Networking and Internet (Globecom 2022 NGNI), Rio de Janeiro, Brazil
3. Nguyen, H. N., S. Souihi, H. A. Tran et S. Fowler. 2022a, «A blockchain-based SDN East/West interface», dans 2022 IEEE Global Communications Conference : Next-Generation Networking and Internet (Globecom 2022 NGNI), Rio de Janeiro, Brazil.
4. Nguyen, T. A. M., S. Bannour Souihi, H. A. Tran et S. Souihi. 2022b, «When NLP meets SDN : an application to global internet exchange network», dans 2022 IEEE International Conference on Communications (ICC) : Next-Generation Networking and Internet Symposium (IEEE ICC'22 - NGNI Symposium), Seoul, Korea (South).
5. Tong, V., Souihi, S., Tran, H. A. and Mellouk, A. (2021). "Machine Learning based Root Cause Analysis for SDN Network". 2021 IEEE Global Communications Conference (GLOBECOM), 1–6. IEEE.
6. D. H. Le, H. A. Tran, S. Souihi and A. Mellouk. "An AI-based Traffic Matrix Prediction Solution for Software-Defined Network". IEEE International Conference on Communications (ICC) 2021
7. D. H. Le , H. A. Tran and S. Souihi. A Reinforcement Learning-Based Solution for Intra-Domain Egress Selection. IEEE International Conference on High Performance Switching and Routing (HPSR 2021).
8. H. A. Tran, Tu Nguyen, S. Souihi, and A. Mellouk. "Towards a novel congestion notification algorithm for a software-defined data center networks.". In IFIP/IEEE International Symposium on Integrated Network Management (IM 2021).
9. V. Tong, S. Souihi, H. A. Tran, and A. Mellouk. "Service-centric Segment Routing Mechanism using Reinforcement Learning for Encrypted Traffic." . In International Conference on Network and Service Management (CNSM 2020) (CNSM 2020).2020.
10. L. Amour, V. Tong, S. Souihi, H. A. Tran and A. Mellouk, "Quality Estimation Framework for Encrypted Traffic (Q2ET)", 2019 IEEE Global Communications Conference (GLOBECOM) 2019, Waikoloa, HI, USA, 2019.
11. T. Chu, V. Tong, H. A. Tran, S. Souihi, D. Tran, and A. Mellouk. 2019. NextLab : A new hybrid testbed and development platform for Software-defined Networking. Tenth International Symposium on Information and Communication Technology (SoICT 2019), Hanoi - Halong Bay, Vietnam, dec, 2019.
12. F. Bannour, S. Souihi and A. Mellouk. "Adaptive Quorum-inspired SLA-Aware Consistency for Distributed ONOS Controllers". 15th International Conference on Network and Service

- Management (CNSM), Halifax, Canada, oct, 2019.
13. V. Tong, H. A. Tran, S. Souihi and A. Mellouk, "A Novel QUIC Traffic Classifier Based on Convolutional Neural Networks", 2018 IEEE Global Communications Conference (GLOBECOM) 2018, Abu Dhabi, United Arab Emirates, 2018.
 14. F. Bannour, S. Souihi and A. Mellouk, "Adaptive State Consistency for Distributed ONOS Controllers", 2018 IEEE Global Communications Conference (GLOBECOM) 2018, Abu Dhabi, United Arab Emirates, 2018.
 15. V. Tong, H. A. Tran, S. Souihi and A. Mellouk, "Empirical study for Dynamic Adaptive Video Streaming Service based on Google Transport QUIC protocol", 2018 IEEE 43rd Conference on Local Computer Networks (LCN), Chicago, IL, USA, 2018.
 16. V. Tong, H. A. Tran, S. Souihi and A. Mellouk, "Network troubleshooting : Survey, Taxonomy and Challenges" 2018 International Conference on Smart Communications in Network Technologies (SaCoNeT), El Oued, 2018.
 17. L. Amour, M. I. Boulabiar, S. Souihi and A. Mellouk, "An improved QoE estimation method based on QoS and affective computing", 2018 International Symposium on Programming and Systems (ISPS), Algiers, 2018.
 18. F. Bannour, S. Souihi and A. Mellouk, "Scalability and reliability aware SDN controller placement strategies", 2017 13th International Conference on Network and Service Management (CNSM), Tokyo, Japan, 2017.
 19. J. Bennaceur, S. Souihi, H. Idoudi, L. A. Saldane and A. Mellouk, "Game-based secure sensing for the mobile cognitive radio network", 2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC), Montreal, QC, Canada 2017.
 20. L. Amour, M. S. Mushtaq, S. Souihi and A. Mellouk, "QoE-Based Framework to Optimize User Perceived Video Quality", 2017 IEEE 42nd Conference on Local Computer Networks (LCN), Singapore, 2017.
 21. L. Amour, S. Sami, M. S. Mushtaq, S. Hoceini and A. Mellouk, "Perceived video quality evaluation based on interactive/repulsive relation between the QoE IFs", 2016 IEEE International Conference on Communications (ICC) 2016, Kuala Lumpur, Malaysia, 2016.
 22. M. A. Senouci, S. Souihi, S. Hoceini and A. Mellouk, "QoE-based network interface selection for heterogeneous wireless networks : A survey and e-Health case proposal", 2016 IEEE Wireless Communications and Networking Conference, Doha, Qatar, 2016.
 23. L. Amour, S. Souihi, S. Hoceini, A. Mellouk, "Building a Large Dataset for Model-based QoE Prediction in the Mobile Environment", In Proceedings of the 18th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM '15). Cancun, Mexico, 2015,
 24. M. Souidi, S. Souihi, S. Hoceini, A. Mellouk. "An adaptive real time mechanism for IaaS cloud provider selection based on QoE aspects". ICC 2015.
 25. S. Souihi, M. Souidi, A. Mellouk. "An Adaptive QoE-Based Network Interface Selection for Multi-homed eHealth Devices". IoT 360 (1) 2015 : 437-442.
 26. L. Amour, S. Souihi, S. Hoceini, A. Mellouk. "An Open Source Platform for Perceived Video

- Quality Evaluation". Q2SWinet@MSWiM 2015 : 139-140.
27. L. Amour, S. Souihi, S. Hoceini, A. Mellouk. A Hierarchical Classification Model of QoE Influence Factors. Wiley WWIC 2015 : 225-238.
 28. S. Souihi, J. Perez, S. Hoceini, A. Mellouk. "A robust, adaptive and hierarchical knowledge dissemination architecture". IEEE Global Communications Conference (GLOBECOM) 2013 : 1489-1494
 29. S. Souihi, S. Hoceini, A. Mellouk, N. A. Saadi. "A hierarchical and multi-criteria knowledge dissemination in autonomic networks". IEEE Global Communications Conference (GLOBECOM) 2012.
 30. S. Souihi, S. Hoceini, A. Mellouk. "A multi-criteria master nodes selection mechanism for knowledge dissemination in autonomic networks". IEEE International Conference on Communications (ICC) 2012, Ottawa, Canada 2012.
 31. S. Souihi, A. Mellouk. "Knowledge Dissemination for Autonomic Network". IEEE International Conference on Communications (ICC) 2011, Kyoto, Japan 2011.

Ouvrages

1. Bannour F, Souihi S. & Mellouk, A. (2022). Software-Defined Networking - vol2. ISTE - Wiley.
2. Tonv V., Souihi S. Tran H. A. & Mellouk, A. (2022). Troubleshooting for Network Operators - The Road to New Paradigm with Encrypted Traffic. ISTE - Wiley. (en cours d'édition)

Bibliographie

Chapitre 1

- Cisco. 2020, «Cisco annual internet report (2018–2023), white paper.», URL "<https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html>". 4
- Clark, D. D., C. Partridge, J. C. Ramming et J. T. Wroclawski. 2003, «A knowledge plane for the internet», dans *Proceedings of the 2003 conference on Applications, technologies, architectures, and protocols for computer communications*, p. 3–10. 4
- Covo, A. A., T. M. Moruzzi et E. D. Peterson. 1989, «Ai-assisted telecommunications network management», dans *1989 IEEE Global Telecommunications Conference and Exhibition 'Communications Technology for the 1990s and Beyond'*, p. 487–491 vol.1. 4
- Feamster, N., J. Rexford et E. Zegura. 2014, «The road to sdn : An intellectual history of programmable networks», *SIGCOMM Comput. Commun. Rev.*, vol. 44, n° 2, doi :10.1145/2602204.2602219, p. 87–98, ISSN 0146-4833. URL <https://doi.org/10.1145/2602204.2602219>. 4
- Leiner, B. M., V. G. Cerf, D. D. Clark, R. E. Kahn, L. Kleinrock, D. C. Lynch, J. Postel, L. G. Roberts et S. Wolff. 2009, «A brief history of the internet», *ACM SIGCOMM Computer Communication Review*, vol. 39, n° 5, p. 22–31. 4
- Qi, J., F. Wu, L. Li et H. Shu. 2007, «Artificial intelligence applications in the telecommunications industry», *Expert Systems*, vol. 24, n° 4, p. 271–291. 4
- Rayward, W. B. 2012, «Paul otlet, an encounter», *Cahiers de la documentation–Bladen voor documentatie*, vol. 2, p. 71. 4
- Robertazzi, T. G. et L. Shi. 2020, «Machine learning in networking», dans *Networking and Computation*, Springer, p. 151–190. 4
- Ross, M., A. Covo et C. Hart. 1988, «An ai-based network management system», dans *IEEE International Performance Computing and Communications Conference*, IEEE Computer Society, Los Alamitos, CA, USA, p. 458,459,460,461, doi :10.1109/PCCC.1988.10121. URL <https://doi.ieeeecomputersociety.org/10.1109/PCCC.1988.10121>. 4
- Sunny, A., R. El-Azouzi, A. Arfaoui, E. Altman, S. Poojary, D. Tsilimantos et S. Valentin. 2019, «Enforcing bitrate-stability for adaptive streaming traffic in cellular networks», *IEEE Transactions on Network and Service Management*, vol. 16, n° 4, p. 1812–1825. 4

Chapitre 2

- Amour, L., M. I. Boulabiar, S. Souihi et A. Mellouk. 2018, «An improved qoe estimation method based on qos and affective computing», dans *2018 International Symposium on Programming and Systems (ISPS)*, IEEE, p. 1–6. [10](#)
- Amour, L., M. S. Mushtaq, S. Souihi et A. Mellouk. 2017, «Qoe-based framework to optimize user perceived video quality», dans *2017 IEEE 42nd Conference on Local Computer Networks (LCN)*, IEEE, p. 599–602. [10](#)
- Amour, L., S. Sami, S. Hoceini et A. Mellouk. 2015a, «Building a large dataset for model-based qoe prediction in the mobile environment», dans *Proceedings of the 18th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, p. 313–317. [10](#)
- Amour, L., S. Sami, S. Hoceini et A. Mellouk. 2015b, «An open source platform for perceived video quality evaluation», dans *Proceedings of the 11th ACM Symposium on QoS and Security for Wireless and Mobile Networks*, p. 139–140. [10](#)
- Amour, L., S. Sami, M. S. Mushtaq, S. Hoceini et A. Mellouk. 2016, «Perceived video quality evaluation based on interactive/repulsive relation between the qoe ifs», dans *2016 IEEE International Conference on Communications (ICC)*, IEEE, p. 1–7. [10](#)
- Amour, L., S. Souihi, S. Hoceini et A. Mellouk. 2015c, «A hierarchical classification model of qoe influence factors», dans *International Conference on Wired/Wireless Internet Communication*, Springer, Cham, p. 225–238. [10](#)
- Amour, L., S. Souihi, A. Mellouk et S. Mushtaq. 2020, «Q2abr : Qoe-aware adaptive video bit rate solution», *International Journal of Communication Systems*, vol. 33, n° 10, p. e4204. [10](#)
- Amour, L., V. Tong, S. Souihi, H. A. Tran et A. Mellouk. 2019, «Quality estimation framework for encrypted traffic (q2et)», dans *2019 IEEE Global Communications Conference (GLOBECOM)*, IEEE, p. 1–6. [10](#), [11](#)
- Balachandran, T., T. Abreu, M. Naloufi, S. Souihi, F. Lucas et A. Janne. 2022, «IoT and transfer learning based urban river quality prediction», dans *2022 IEEE Global Communications Conference : IoT and Sensor Networks (Globecom 2022 IoTSN)*, Rio de Janeiro, Brazil. [12](#)
- Bannour, F., S. Souihi et A. Mellouk. 2017a, «Distributed sdn control : Survey, taxonomy, and challenges», *IEEE Communications Surveys & Tutorials*, vol. 20, n° 1, p. 333–354. [11](#)
- Bannour, F., S. Souihi et A. Mellouk. 2017b, «Scalability and reliability aware sdn controller placement strategies», dans *2017 13th International Conference on Network and Service Management (CNSM)*, IEEE, p. 1–4. [11](#)
- Bannour, F., S. Souihi et A. Mellouk. 2018, «Adaptive state consistency for distributed onos controllers», dans *2018 IEEE Global Communications Conference (Globecom)*, IEEE, p. 1–7. [11](#)
- Bannour, F., S. Souihi et A. Mellouk. 2019, «Adaptive quorum-inspired sla-aware consistency for distributed sdn controllers», dans *2019 15th International Conference on Network and Service Management (CNSM)*, IEEE, p. 1–7. [11](#)
- Bannour, F., S. Souihi et A. Mellouk. 2020, «Adaptive distributed sdn controllers : application to content-centric delivery networks», *Future Generation Computer Systems*, vol. 113, p. 78–93. [11](#)
- Bennaceur, J., S. Souihi, H. Idoudi, L. A. Saldane et A. Mellouk. 2017, «Game-based secure sensing for the mobile cognitive radio network», dans *2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, IEEE, p. 1–7. [11](#)

- Chu, T.-T., V. Tong, H. A. Tran, S. Souihi, D. Q. Tran et A. Mellouk. 2019, «Nextlab : A new hybrid testbed and development platform for software-defined networking», dans *Proceedings of the Tenth International Symposium on Information and Communication Technology*, p. 186–190. [11](#), [12](#)
- Do, Q. H., S. Souihi, V. Tong, H. A. Tran et S. Mhadhbi. 2022, «How tezos blockchain can meet IoT?», dans *2022 IEEE Global Communications Conference : Next-Generation Networking and Internet (Globecom 2022 NGNI)*, Rio de Janeiro, Brazil. [12](#)
- Hoang, N.-T., H.-N. Nguyen, H.-A. Tran et S. Souihi. 2022, «A novel adaptive east&west interface for a heterogeneous and distributed sdn network», *Electronics*, vol. 11, n° 7, doi :10.3390/electronics11070975, ISSN 2079-9292. URL <https://www.mdpi.com/2079-9292/11/7/975>. [12](#)
- Nam Nguyen, H., H. Anh Tran, S. Fowler et S. Souihi. 2021, «A survey of blockchain technologies applied to software-defined networking : Research challenges and solutions», *IET Wireless Sensor Systems*, vol. 11, n° 6, doi :https://doi.org/10.1049/wss2.12031, p. 233–247. URL <https://ietresearch.onlinelibrary.wiley.com/doi/abs/10.1049/wss2.12031>. [12](#)
- Nguyen, H. N., S. Souihi, H. A. Tran et S. Fowler. 2022a, «A blockchain-based SDN East/West interface», dans *2022 IEEE Global Communications Conference : Next-Generation Networking and Internet (Globecom 2022 NGNI)*, Rio de Janeiro, Brazil. [12](#)
- Nguyen, T. A. M., S. Bannour Souihi, H. A. Tran et S. Souihi. 2022b, «When NLP meets SDN : an application to global internet exchange network», dans *2022 IEEE International Conference on Communications (ICC) : Next-Generation Networking and Internet Symposium (IEEE ICC'22 - NGNI Symposium)*, Seoul, Korea (South). [12](#)
- Tong, V., S. Souihi, H. A. Tran et A. Mellouk. 2021a, «Machine learning based root cause analysis for sdn network», dans *2021 IEEE Global Communications Conference (GLOBECOM)*, IEEE, p. 1–6. [11](#)
- Tong, V., S. Souihi, H. A. Tran et A. Mellouk. 2021b, «Sdn-based application-aware segment routing for large-scale network», *IEEE Systems Journal*. [11](#)
- Tong, V., H. A. Tran, S. Souihi et A. Mellouk. 2018, «A novel quic traffic classifier based on convolutional neural networks», dans *2018 IEEE Global Communications Conference (GLOBECOM)*, IEEE, p. 1–6. [11](#)
- Van, T., S. Souihi, H. A. Tran et A. Mellouk. 2020, «Service-centric segment routing mechanism using reinforcement learning for encrypted traffic», dans *2020 16th International Conference on Network and Service Management (CNSM)*, IEEE, p. 1–5. [11](#)
- Van, T., H. A. Tran, S. Souihi et A. Mellouk. 2018a, «Empirical study for dynamic adaptive video streaming service based on google transport quic protocol», dans *2018 IEEE 43rd Conference on Local Computer Networks (LCN)*, IEEE, p. 343–350. [11](#)
- Van, T., H. A. Tran, S. Souihi et A. MELLOUK. 2018b, «Network troubleshooting : survey, taxonomy and challenges», dans *2018 International Conference on Smart Communications in Network Technologies (SaCoNeT)*, IEEE, p. 165–170. [11](#)

Chapitre 3

- «Pox», <http://www.noxrepo.org/pox/about-pox/>. Accessed : 2015-11-22. 26
- 2015, «Floodlight Project», URL <http://www.projectfloodlight.org/>, accessed : 2015-12-07. 26
- 2016, «Opendaylight project», URL <http://www.opendaylight.org/>, accessed : 2016.-01-05. 26, 27, 35
- ONOS. URL <https://onosproject.org/>. 35
- Ahmadi, V., A. Jalili, S. M. Khorramizadeh et M. Keshtgari. 2015, «A hybrid nsga-ii for solving multiobjective controller placement in sdn», dans *The 2nd International Conference on Knowledge-Based Engineering and Innovation (KBEI)*, p. 663–669, doi :10.1109/KBEI.2015.7436122. 29
- Aslan, M. et A. Matrawy. 2016, «Adaptive consistency for distributed SDN controllers», dans *2016 17th International Telecommunications Network Strategy and Planning Symposium (Networks)*, p. 150–157, doi :10.1109/NETWKS.2016.7751168. 35
- Bailis, P., S. Venkataraman, M. J. Franklin, J. M. Hellerstein et I. Stoica. 2012, «Probabilistically bounded staleness for practical partial quorums», *CoRR*, vol. abs/1204.6082. URL <http://arxiv.org/abs/1204.6082>. 39, 43
- Bailis, P., S. Venkataraman, M. J. Franklin, J. M. Hellerstein et I. Stoica. 2014, «Quantifying eventual consistency with pbs», *The VLDB Journal*, vol. 23, n° 2, p. 279–302. 43
- Bannour, F., S. Souihi et A. Mellouk. 2017, «Scalability and Reliability Aware SDN Controller Placement Strategies», dans *Proc. Of the 13th International Conf. on Network and Service Management, CNSM 2017, MANSDNNFV2017*, Tokyo, Japan. URL <https://hal.archives-ouvertes.fr/hal-01676567>. 30
- Bannour, F., S. Souihi et A. Mellouk. 2018, «Distributed sdn control : Survey, taxonomy, and challenges», *IEEE Communications Surveys Tutorials*, vol. 20, n° 1, doi :10.1109/COMST.2017.2782482, p. 333–354. 26
- Berde, P., M. Gerola, J. Hart, Y. Higuchi, M. Kobayashi, T. Koide, B. Lantz, B. O’Connor, P. Radoslavov, W. Snow et G. Parulkar. 2014, «Onos : Towards an open, distributed SDN OS», dans *Proceedings of the Third Workshop on Hot Topics in Software Defined Networking, HotSDN ’14*, ACM, New York, NY, USA, ISBN 978-1-4503-2989-7, p. 1–6, doi :10.1145/2620728.2620744. 26, 27
- Botelho, F., A. Bessani, F. M. V. Ramos et P. Ferreira. 2014, «On the design of practical fault-tolerant sdn controllers», dans *2014 Third European Workshop on Software Defined Networks*, ISSN 2379-0350, p. 73–78, doi :10.1109/EWSDN.2014.25. 26
- Brakmo, L. S., S. W. O’Malley et L. L. Peterson. 1994, «Tcp vegas : New techniques for congestion detection and avoidance», vol. 24, n° 4, doi :10.1145/190809.190317, p. 24–35, ISSN 0146-4833. URL <https://doi.org/10.1145/190809.190317>. 41
- Clark, D. D., C. Partridge, J. C. Ramming et J. T. Wroclawski. 2003, «A knowledge plane for the internet», dans *Proceedings of the 2003 conference on Applications, technologies, architectures, and protocols for computer communications*, p. 3–10. 24
- Curtis, A. R., J. C. Mogul, J. Tourrilhes, P. Yalagandula, P. Sharma et S. Banerjee. 2011, «Devoflow : Scaling flow management for high-performance networks», dans *Proceedings of the ACM SIGCOMM 2011 Conference, SIGCOMM ’11*, ACM, New York, NY, USA, ISBN 978-1-4503-0797-0, p. 254–265, doi :10.1145/2018436.2018466. 26

- Das, T., V. Sridharan et M. Gurusamy. 2019, «A survey on controller placement in SDN», *IEEE Communications Surveys Tutorials*, doi :10.1109/COMST.2019.2935453, p. 1–1. [27](#)
- Feamster, N., J. Rexford et E. Zegura. 2014, «The road to sdn : An intellectual history of programmable networks», *SIGCOMM Comput. Commun. Rev.*, vol. 44, n° 2, doi :10.1145/2602204.2602219, p. 87–98, ISSN 0146-4833. URL <https://doi.org/10.1145/2602204.2602219>. [24](#)
- Ganatra, N. S. A. 2012, «Comparative study of several clustering algorithms.», *International Journal of Advanced Computer Research*, p. 37–42. [29](#)
- Gomes, E. R. et R. Kowalczyk. 2009, «Dynamic analysis of multiagent q-learning with epsilon-greedy exploration», dans *Proceedings of the 26th International Conference on Machine Learning*. [41](#)
- Gude, N., T. Koponen, J. Pettit, B. Pfaff, M. Casado, N. McKeown et S. Shenker. 2008, «Nox : Towards an operating system for networks», *SIGCOMM Comput. Commun. Rev.*, vol. 38, n° 3, doi :10.1145/1384609.1384625, p. 105–110, ISSN 0146-4833. [26](#)
- Gupta, A., M. Shahbaz, L. Vanbever, H. Kim, R. Clark, N. Feamster, J. Rexford et S. Shenker. 2014, «Sdx : A software defined internet exchange», *ACM SIGCOMM*. [26](#)
- Hassas Yeganeh, S. et Y. Ganjali. 2012, «Kandoo : A framework for efficient and scalable offloading of control applications», dans *Proceedings of the First Workshop on Hot Topics in Software Defined Networks, HotSDN '12*, ACM, New York, NY, USA, ISBN 978-1-4503-1477-0, p. 19–24, doi :10.1145/2342441.2342446. [26](#)
- Haßlinger, G. et O. Hohlfeld. 2008, «The gilbert-elliott model for packet loss in real time services on the internet», dans *14th GI/ITG Conference-Measurement, Modelling and Evaluation of Computer and Communication Systems*, VDE, p. 1–15. [42](#)
- Heller, B., R. Sherwood et N. McKeown. 2012, «The controller placement problem», dans *Proceedings of the First Workshop on Hot Topics in Software Defined Networks, HotSDN '12*, ACM, New York, NY, USA, ISBN 978-1-4503-1477-0, p. 7–12, doi :10.1145/2342441.2342444. [28](#)
- Hock, D., M. Hartmann, S. Gebert, M. Jarschel, T. Zinner et P. Tran-Gia. 2013, «Pareto-optimal resilient controller placement in sdn-based core networks», dans *25th International Teletraffic Congress (ITC)*, Shanghai, China. [30](#)
- Jain, S., A. Kumar, S. Mandal, J. Ong, L. Poutievski, A. Singh, S. Venkata, J. Wanderer, J. Zhou, M. Zhu, J. Zolla, U. Hölzle, S. Stuart et A. Vahdat. 2013, «B4 : Experience with a globally-deployed software defined wan», dans *Proceedings of the ACM SIGCOMM 2013 Conference on SIGCOMM*, SIGCOMM '13, ACM, New York, NY, USA, ISBN 978-1-4503-2056-6, p. 3–14, doi : 10.1145/2486001.2486019. [26](#)
- Katta, N., H. Zhang, M. Freedman et J. Rexford. 2015, «Ravana : Controller fault-tolerance in software-defined networking», dans *Proceedings of the 1st ACM SIGCOMM Symposium on Software Defined Networking Research, SOSR '15*, ISBN 978-1-4503-3451-8, p. 4 :1–4 :12. [26](#)
- Killi, B. P. R. et S. V. Rao. 2019, «Controller placement in software defined networks : A comprehensive survey», *Computer Networks*, vol. 163, p. 106 883, ISSN 1389-1286. [28](#)
- Klassen, F. *TcpReplay*. <https://tcpreplay.appneta.com>. [42](#)
- Koponen, T., M. Casado, N. Gude, J. Stribling, L. Poutievski, M. Z. Google, R. Ramanathan, Y. I. NEC, H. I. NEC, T. H. NEC et S. Shenker. 2010, «Onix : a distributed control platform for large-scale production networks», dans *9th Conference on Operating Systems Design and Implementation*, p. 351–364. [26](#), [27](#)

- Lakshman, A. et P. Malik. 2010, «Cassandra : A decentralized structured storage system», *SIGOPS Oper. Syst. Rev.*, vol. 44, n° 2, doi :10.1145/1773912.1773922, p. 35–40, ISSN 0163-5980. 35
- Lange, S., S. Gebert, T. Zinner, P. Tran-Gia, D. Hock, M. Jarschel et M. Hoffmann. 2015, «Heuristic approaches to the controller placement problem in large scale sdn networks.», *IEEE Transactions on Network and Service Management*, vol. 12, n° 1, p. 4–17. 28
- Liang, K., L. Zhao, X. Chu et H. H. Chen. 2017, «An integrated architecture for software defined and virtualized radio access networks with fog computing», *IEEE Network*, vol. 31, n° 1, doi : 10.1109/MNET.2017.1600027NM, p. 80–87, ISSN 0890-8044. 46
- Ojo, M., D. Adami et S. Giordano. 2016, «A SDN-IoT architecture with NFV implementation», dans *2016 IEEE Globecom Workshops (GC Wkshps)*, p. 1–6, doi :10.1109/GLOCOMW.2016.7848825. 46
- ONF. «Open Networking Foundation», <https://www.opennetworking.org/>. Accessed : 2016-05-19. 24
- Ongaro, D. et J. Ousterhout. 2014, «In search of an understandable consensus algorithm», dans *2014 USENIX Annual Technical Conference (USENIX ATC 14)*. 35
- Phemius, K., M. Bouet et J. Leguay. 2013, «DISCO : distributed multi-domain SDN controllers», *CoRR*, vol. abs/1308.6138. 26
- Shraddha, K. et M. Emmanuel. 2014, «Review and comparative study of clustering techniques.», *International Journal of Computer Science and Information Technology (IJCSIT)*, vol. 5, p. 805–812. 33
- Sivasubramanian, S. 2012, «Amazon dynamodb : A seamlessly scalable non-relational database service», dans *Proceedings of the 2012 ACM SIGMOD International Conference on Management of Data*, SIGMOD '12, ACM, New York, NY, USA, ISBN 978-1-4503-1247-9, p. 729–730, doi :10.1145/2213836.2213945. 35, 40
- Sutton, R. S. et A. G. Barto. 2018, *Reinforcement Learning : An Introduction*, 2^e éd., The MIT Press. URL <http://incompleteideas.net/book/the-book-2nd.html>. 40
- Tootoonchian, A. et Y. Ganjali. 2010, «Hyperflow : A distributed control plane for openflow», dans *Proceedings of the 2010 Internet Network Management Conference on Research on Enterprise Networking*, INM/WREN'10, Berkeley, CA, USA, p. 3–3. 26, 27
- Watkins, C. J. C. H. et P. Dayan. 1992, «Q-learning», dans *Machine Learning*, p. 279–292. 40
- Yu, H. et A. Vahdat. 2000, «Design and evaluation of a continuous consistency model for replicated services», dans *Proceedings of the 4th Conference on Symposium on Operating System Design & Implementation - Volume 4*, OSDI'00, Berkeley, CA, USA. 35, 36
- Yu, M., J. Rexford, M. J. Freedman et J. Wang. 2010, «Scalable flow-based networking with difane», dans *Proceedings of the ACM SIGCOMM 2010 Conference*, SIGCOMM '10, ACM, New York, NY, USA, ISBN 978-1-4503-0201-2, p. 351–362, doi :10.1145/1851182.1851224. 26

Chapitre 4

2005. Access : www.vqeq.org, «Video Quality Experts Group, Multimedia Group Test Plan, Draft version 1.8», . 62

- 2016, «International Telecommunication Union - Telecommunication (ITU-T)-T Recommendation P.800.1 : Methods for Objective and Subjective Assessment of Speech and Video Quality : Mean Opinion Score (MOS) Terminology», , p. 1–18. [61](#)
- November 2009, «Human Factors (HF) : Quality of Experience (QoE) Requirements for Real-time Communication Services», dans *ETSI TR 102 643*, p. 1–37. [61](#)
- pages 1-8. January 2006, «ITU-T Recommendation P.10/G.100 : Definition of Quality of Experience (QoE)», . [61](#)
- A. Appriou. November 1999, «Multisensor Signal Processing in the Framework of the Theory of Evidence», *In Application of Mathematical Signal Processing Techniques to Mission Systems, Research and Technology Organization*, p. 1–31. [66](#)
- A. Khan, L. Sun and E. Ifeachor. 2010, «Learning Models for Video Quality Prediction over Wireless Local Area Network and Universal Mobile Telecommunication System Networks», *IET Communications*, vol. 4, n° 12, p. pages 1389–1403, ISSN 1751-8628. [62](#)
- A. Martin, A. L. Jusselme and C. Osswald. 2008, «Conflict Measure for the Discounting Operation on Belief Functions», dans *11th International Conference on Information Fusion*, p. 1–8. [65](#)
- A. P. Dempster. 325-339. February 1967, «Upper and Lower Probabilities induced by a Multivalued Mapping», *Annals of Mathematical Statistics Journal*, vol. Volume 38, n° 2. [65](#)
- Adami, D., G. Antichi, R. G. Garroppo, S. Giordano et A. W. Moore. 2015, «Towards an sdn network control application for differentiated traffic routing», dans *2015 IEEE International Conference on Communications (ICC)*, IEEE, p. 5827–5832. [52](#), [53](#)
- Alreshoodi, M. et J. Woods. June 2014, «Survey on QoE\QoS Correlation Models For Multimedia Services», *Computing Research Repository (CoRR) Journal*, vol. Volume 1306.0221, p. pages 53–72. [61](#), [67](#)
- Amour, L., S. Souihi et A. Mellouk. October 2017, «Dataset 01 : Controlled Laboratory dataset collecting You Tube QoE IFs using VLC video player. Access : », <https://github.com/Lamyne/QoE-Dataset-VLC-Indicators/>. [63](#)
- Aroussi, S. et A. Mellouk. 2014, «Survey on Machine Learning-based QoE-QoS correlation models», dans *International Conference on Computing, Management and Telecommunications (ComManTel)*, p. 200–204. [61](#), [67](#)
- Aydin, B., A. A. Kindiroglu, O. Aran et L. Akarun. 2016, «Automatic Personality Prediction from Audiovisual Data using Random Forest Regression», dans *2016 23rd International Conference on Pattern Recognition (ICPR)*, p. 37–42. [62](#)
- P. Le Callet, S. M. et A. Perkis. March 2013, «Qualinet White Paper on Definitions of Quality of Experience (QoE) and Related Concepts», dans *European Network on Quality of Experience in Multimedia Systems and Services*, p. 1–25. [61](#)
- Chen, H., X. Yu et L. Xie. May 2013, «End-to-end Quality Adaptation Scheme based on QoE Prediction for Video Streaming Service in LTE networks», dans *2013 11th International Symposium and Workshops on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt)*, p. 627–633. [62](#)
- CholletKeras, F. 2016, «loss function», <https://keras.io/losses/>. [57](#)
- C.K. Murphy. July 2000, «Combining Belief Functions when Evidence Conflicts», *Decision Support System Journal*, vol. Volume 29, n° 1, p. pages 1–9. [66](#)

- E. Lefevre. June 2002, «Belief function combination and conflict management», *Information Fusion Journal*, vol. Volume 3(2), p. pages 149–162. 66
- Fan, Z. et R. Liu. 2017, «Investigation of machine learning based network traffic classification», dans *2017 International Symposium on Wireless Communication Systems (ISWCS)*, IEEE, p. 1–6. 53
- Fathi-Kazerooni, S., Y. Kaymak et R. Rojas-Cessa. 2019, «Tracking user application activity by using machine learning techniques on network traffic», dans *2019 International Conference on Artificial Intelligence in Information and Communication (ICAIIIC)*, IEEE, p. 405–410. 53
- Florea, M. C., J. Dezert, P. Valin, F. Smarandache et A. Jousselme. April 2006, «Adaptative Combination Rule and Proportional Conflict Redistribution Rule for Information Fusion», *Computing Research Repository (CoRR) Journal*, p. 1–8. 66
- Freund, Y. et R. E. Schapire. pages 148–156. July 1996, «Experiments with a new boosting algorithm», dans *Proceedings of the Thirteenth International Conference On Machine Learning*, Morgan Kaufmann Publisher. 62
- G. Shafer. Princeton University Press, pages 1-314. April 1976, *A Mathematical Theory of Evidence*. 64, 66
- Goodfellow, I., Y. Bengio et A. Courville. 2016, *Deep Learning*, MIT Press. <http://www.deeplearningbook.org>. 54
- H. Mark and F. Eibe and H. Geoffrey and P. Bernhard and R. Peter and W. Ian. June 2009, «The WEKA Data Mining Software : An Update», *ACM SIGKDD explorations newsletter*, , n° 1, p. pages 10–18. 61
- He, K. et J. Sun. 2015, «Convolutional neural networks at constrained time cost», dans *Proceedings of the IEEE conference on computer vision and pattern recognition*, p. 5353–5360. 56
- Ickin, S., K. Wac, M. Fiedler, L. Janowski, J. H. Hong et A. K. Dey. 2012, «Factors Influencing Quality of Experience of commonly used Mobile Applications», *IEEE Communications Magazine*, vol. 50, n° 4, doi:10.1109/MCOM.2012.6178833, p. 48–56, ISSN 0163-6804. 61
- J. Mendes-Moreira, J. C. M. A., Soares et J. Sousa. 2012, «Ensemble Approaches for Regression : A Survey», *ACM Comput. Surv.*, vol. 45, n° 1, p. 10 :1–10 :40, ISSN 0360-0300. 62
- Khan, A., L. Sun et E. Ifeachor. 2009, «Content-Based Video Quality Prediction for MPEG4 Video Streaming over Wireless Networks», *Journal of Multimedia*, vol. 4, n° 4, p. pages 228–239. 62
- L. Amour, S. Souihi, M. S. Mushtaq, S. Hoceini and A. Mellouk. 2016, «Perceived video quality evaluation based on interactive/repulsive relation between the QoE IFs», dans *2016 IEEE International Conference on Communications (ICC)*, p. 1–7. 62
- Langley, A., A. Riddoch, A. Wilk, A. Vicente, C. Krasic, D. Zhang, F. Yang, F. Kouranov, I. Swett, J. Iyengar et collab.. 2017, «The quic transport protocol : Design and internet-scale deployment», dans *Proceedings of the Conference of the ACM Special Interest Group on Data Communication*, ACM, p. 183–196. 52
- Liotou, E., D. Tsolkas et N. Passas. 2016, «A Roadmap on QoE Metrics and Models», dans *2016 23rd International Conference on Telecommunications (ICT)*, p. 1–5. 61
- Lopez-Martin, M., B. Carro, A. Sanchez-Esguevillas et J. Lloret. 2017, «Network traffic classifier with convolutional and recurrent neural networks for internet of things», *IEEE Access*, vol. 5, p. 18 042–18 050. 52, 53, 54

- Lotfollahi, M., M. J. Siavoshani, R. S. H. Zade et M. Saberian. 2020, «Deep packet : A novel approach for encrypted traffic classification using deep learning», *Soft Computing*, vol. 24, n° 3, p. 1999–2012. [53](#)
- M. Galar and A. Fernandez and E. Barrenechea and H. Bustince and F. Herrera. 2012, «A Review on Ensembles for the Class Imbalance Problem : Bagging, Boosting, and Hybrid-Based Approaches», *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 42, n° 4, p. pages 463–484. [62](#)
- Mushtaq, M. S., S. Fowler, B. Augustin et A. Mellouk. 2016, «QoE in 5G Cloud Networks using Multimedia services», dans *IEEE Wireless Communications and Networking Conference*, p. 1–6. [61](#)
- P. K. Mallapragada and R. Jin and A. K. Jain and Y. Liu. 2009, «SemiBoost : Boosting for Semi-Supervised Learning», *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, n° 11, p. pages 2000–2014. [62](#)
- P. Walley. pages 1-700. December 1991, *Statistical Reasoning with Imprecise Probabilities*, Chapman & Hall/CRC Monographs on Statistics & Applied Probability, Taylor & Francis Publisher, ISBN 9780412286605. [65](#)
- Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg et collab.. 2011, «Scikit-learn : Machine learning in python», *Journal of machine learning research*, vol. 12, n° Oct, p. 2825–2830. [57](#)
- Reichl, P., S. Egger, R. Schatz et A. D’Alconzo. 2010, «The Logarithmic Nature of QoE and the Role of the Weber-Fechner Law in QoE Assessment», dans *2010 IEEE International Conference on Communications*, ISSN 1550-3607, p. 1–5. [61](#)
- S. Le Hegarat-Masclé, I. Bloch and D. Vidal-Madjar. 1997, «Application of Dempster-Shafer evidence theory to unsupervised classification in multisource remote sensing», *IEEE Transactions on Geoscience and Remote Sensing*, vol. 35, n° 4, p. pages 1018–1031. [65](#)
- Salman, O., I. H. Elhajj, A. Chehab et A. Kayssi. 2018, «A multi-level internet traffic classifier using deep learning», dans *2018 9th International Conference on the Network of the Future (NOF)*, IEEE, p. 68–75. [53](#)
- Smets, P. 1998, *The Transferable Belief Model for Quantified Belief Representation*, Springer Netherlands, Dordrecht, ISBN 978-94-017-1735-9, p. 267–301, doi :10.1007/978-94-017-1735-9_9. URL https://doi.org/10.1007/978-94-017-1735-9_9. [64](#), [65](#), [66](#)
- Smets, P. August 1993, «Belief Functions : The Disjunctive Rule of Combination and the Generalized Bayesian Theorem», *International Journal of Approximate Reasoning*, vol. Volume 9, n° 1, doi :[http://dx.doi.org/10.1016/0888-613X\(93\)90005-X](http://dx.doi.org/10.1016/0888-613X(93)90005-X), p. pages 1–35, ISSN 0888-613X. URL [//www.sciencedirect.com/science/article/pii/0888613X9390005X](http://www.sciencedirect.com/science/article/pii/0888613X9390005X). [66](#)
- Sokolova, M. et G. Lapalme. 2009, «A systematic analysis of performance measures for classification tasks», *Information Processing & Management*, vol. 45, n° 4, p. 427–437. [56](#)
- T. Denoeux. 1995, «A k-Nearest Neighbor Classification Rule based on Dempster-Shafer Theory», *IEEE Transactions on Systems, Man, and Cybernetics*, vol. Volume 25, n° 5, p. pages 804–813, ISSN 0018-9472. [66](#)
- Tantithamthavorn, C., S. McIntosh, A. E. Hassan et K. Matsumoto. 2016, «Automated Parameter Optimization of Classification Techniques for Defect Prediction Models», dans *2016 IEEE/ACM 38th International Conference on Software Engineering (ICSE)*, p. 321–332. [62](#)

- Ujjan, R. M. A., Z. Pervez, K. Dahal, A. K. Bashir, R. Mumtaz et J. González. 2020, «Towards sflow and adaptive polling sampling for deep learning based ddos detection in sdn», *Future Generation Computer Systems*, vol. 111, p. 763–779. 54
- Wang, P., F. Ye, X. Chen et Y. Qian. 2018, «Datanet : Deep learning based encrypted network traffic classification in sdn home gateway», *IEEE Access*, vol. 6, p. 55 380–55 391. 53
- Wang, W., M. Zhu, J. Wang, X. Zeng et Z. Yang. 2017, «End-to-end encrypted traffic classification with one-dimensional convolution neural networks», dans *2017 IEEE International Conference on Intelligence and Security Informatics (ISI)*, IEEE, p. 43–48. 53
- Youssef, Y. B., A. Mellouk, M. Afif et S. Tabbane. December 2016, «Video Quality Assessment Based on Statistical Selection Approach for QoE Factors Dependency», dans *2016 IEEE Global Communications Conference (GLOBECOM)*, p. 1–6. 62
- Zhou, K., W. Wang, C. Wu et T. Hu. 2020, «Practical evaluation of encrypted traffic classification based on a combined method of entropy estimation and neural networks», *ETRI Journal*, vol. 42, n° 3, p. 311–323. 54

Chapitre 5

- February 2021, «Mininet», URL <http://mininet.org/>. 81
- February 2021, «Open network operating system (onos)», URL <https://wiki.onosproject.org/display/ONOS/ONOS>. 81
- July 2016. p. 1-42, «International telecommunication union - telecommunication (itu-t) recommendation p800.1, mean opinion score (mos) terminology.», . 84, 89
- Azumi, M., T. Kurosaka et M. Bandai. 2015, «A qoe-aware quality-level switching algorithm for adaptive video streaming», dans *2015 IEEE Global Communications Conference (GLOBECOM'15)*, p. 1–5, doi :10.1109/GLOCOM.2015.7417622. 84
- Bahnasse, A., F. E. Louhab, H. A. Oulahyane, M. Talea et A. Bakali. 2018, «Novel sdn architecture for smart mpls traffic engineering-diffserv aware management», *Future Generation Computer Systems*, vol. 87, p. 115–126. 78
- Barakabitze, A. A., I.-H. Mkwawa, A. Hines, L. Sun et E. Ifeakor. 2020, «Qoemultisdn : Management of multimedia services using mptcp/sr in softwarized and virtualized networks», *IEEE Access*. 78
- C. Liu, I. B. et M. Gabbouj. February 2011. p. 169-174, «Rate adaptation for adaptive http streaming», dans *Proceedings of the Second Annual ACM Conference on Multimedia Systems (MM-Sys'11)*, ISBN 978-1-4503-0518-1. 84
- CISCO. February 2021, «Cisco white paper report : Encrypted traffic analytics», URL <https://www.cisco.com/c/dam/en/us/solutions/collateral/enterprise-networks/enterprise-network-security/nb-09-encrytd-traf-anlytcs-wp-cte-en.pdf>. 77
- cisco. January 2021a, «Cisco and vodafone showcase mobile transport networking advancements via segment routing at mobile world congress», . 77
- cisco. January 2021b, «Cisco to optimize ntt docomo's 5g mobile backhaul for simpler, more flexible and scalable network operation», . 77
- CISCO Systems, I. June 2017, «Cisco white paper : Visual networking index : Forecast and methodology, 2016-2021», , p. 1–17. 83

- Clark, D. D., C. Partridge et J. C. Ramming. 2003, «A knowledge plane for the internet», dans *In SIGCOMM*, p. 3–10. [76](#)
- Custura, A., R. Secchi et G. Fairhurst. 2018, «Exploring dscp modification pathologies in the internet», *Computer Communications*, vol. 127, p. 86–94. [77](#), [78](#)
- Dong, C., C. Zhang, Z. Lu, B. Liu et B. Jiang. 2020, «Cetalytics : Comprehensive effective traffic information analytics for encrypted traffic classification», *Computer Networks*, vol. 176, p. 107258. [77](#)
- Feamster, N., J. Rexford et E. Zegura. 2014, «The road to sdn : An intellectual history of programmable networks», *SIGCOMM Comput. Commun. Rev.*, vol. 44, n° 2, doi :10.1145/2602204.2602219, p. 87–98, ISSN 0146-4833. URL <https://doi.org/10.1145/2602204.2602219>. [77](#)
- Grigoriou, E., A. A. Barakabitze, L. Atzori, L. Sun et V. Piloni. 2017, «An sdn-approach for qoe management of multimedia services using resource allocation», dans *2017 IEEE International Conference on Communications (ICC)*, IEEE, p. 1–7. [82](#)
- Han, B., F. Qian, L. Ji et V. Gopalakrishnan. 2016, «Mp-dash : Adaptive video streaming over preference-aware multipath», dans *Proceedings of the 12th International on Conference on Emerging Networking EXperiments and Technologies*, (CoNEXT'16), ACM, ISBN 978-1-4503-4292-6, p. 129–143. [86](#), [87](#)
- He, Y.-L., X.-L. Zhang, W. Ao et J. Z. Huang. 2018, «Determining the optimal temperature parameter for softmax function in reinforcement learning», *Applied Soft Computing*, vol. 70, p. 80–85. [81](#)
- J. Jiang, V. S. et H. Zhang. January 2014, «Improving fairness, efficiency, and stability in http-based adaptive video streaming with festive», dans *Proceedings of the 8th International Conference on Emerging Networking Experiments and Technologies*, (CoNEXT'12), ISBN 978-1-4503-1775-7, p. 97–108. [84](#)
- Jin, C., Z. Yang, Z. Wang et M. I. Jordan. 2020, «Provably efficient reinforcement learning with linear function approximation», dans *Conference on Learning Theory*, PMLR, p. 2137–2143. [80](#)
- Joseph, V. et G. de Veciana. 2014, «Nova : Qoe-driven optimization of dash-based video delivery in networks», dans *IEEE INFOCOM 2014 - IEEE Conference on Computer Communications*, p. 82–90. [84](#)
- Juluri, P., V. Tamarapalli et D. Medhi. 2015, «Sara : Segment aware rate adaptation algorithm for dynamic adaptive streaming over http», dans *2015 IEEE International Conference on Communication Workshop (ICCW)*, p. 1765–1770. [84](#)
- Koenig, S. et R. G. Simmons. 1993, «Complexity analysis of real-time reinforcement learning», dans *AAAI*, p. 99–107. [80](#)
- Kukreja, N., R. Alvizu, A. Kos, G. Maier, R. Morro, A. Capello et C. Cavazzoni. 2016, «Demonstration of sdn-based orchestration for multi-domain segment routing networks», dans *2016 18th International Conference on Transparent Optical Networks (ICTON)*, IEEE, p. 1–4. [77](#), [78](#)
- Kushwaha, A., S. Sharma, N. Bazard, A. Gumaste et B. Mukherjee. 2020, «Design, analysis, and a terabit implementation of a source-routing-based sdn data plane», *IEEE Systems Journal*. [77](#)
- Li, Z., S. Peng, D. Voyer, C. Xie, P. Liu, Z. Qin, K. Ebisawa, S. Previdi et J. Guichard. 2020, «Problem statement and use cases of application-aware networking (apn)», Internet-Draft draft-li-apn-problem-statement-usecases-01, Internet Engineering Task Force. URL <https://datatracker.ietf.org/doc/html/draft-li-apn-problem-statement-usecases-01>, work in Progress. [77](#), [78](#)

- Lin, S.-C., I. F. Akyildiz, P. Wang et M. Luo. 2016, «Qos-aware adaptive routing in multi-layer hierarchical software defined networks : A reinforcement learning approach», dans *2016 IEEE International Conference on Services Computing (SCC)*, IEEE, p. 25–33. 82
- Liotou, E., D. Tsolkas et N. Passas. 2016, «A roadmap on QoE metrics and models», dans *2016 23rd International Conference on Telecommunications (ICT'16)*, p. 1–5. 84
- M. Claeys, S. L., J. Famaey, T. Wu, L. Van et F. D. Turck. May 2013, «Design of a q-learning-based client quality selection algorithm for http adaptive video streaming», dans *Adaptive and Learning Agents Workshop*, p. 30–37. 84
- Mammeri, Z. 2019, «Reinforcement learning based routing in networks : Review and classification of approaches», *IEEE Access*, vol. 7, p. 55 916–55 950. 80
- McLean, P. 2009, «Apple launches http live streaming standard in iphone 3.0», *AppleInsider*. Retrieved 15 October 2011. 84, 88
- Menkovski, V. et A. Liotta. 2013, «Intelligent control for adaptive video streaming», dans *2013 IEEE International Conference on Consumer Electronics (ICCE'13)*, p. 127–128. 84, 85
- Mushtaq, M., B. Augustin et A. Mellouk. 2015, «Regulating QoE for adaptive video streaming using BBF method», dans *2015 IEEE International Conference on Communications (ICC'15)*, p. 6855–6860. 84
- Peng, S., J. Mao, R. Hu et Z. Li. 2020, «Demo abstract : Apn6 : Application-aware ipv6 networking», dans *IEEE INFOCOM 2020-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPs)*, IEEE, p. 1330–1331. 77, 78
- Priyadarshana, W. J. et G. Sofronov. 2015, «Multiple break-points detection in array cgh data via the cross-entropy method», *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 12, n° 2, p. 487–498, ISSN 1545-5963. 85, 86
- R. Bertrand, P., M. Fhima et A. Guillin. 2010. p 1-26, «Off-line detection of multiple change points with the filtered derivative with p-value method», *Sequential Analysis Journal*. 84
- Rego, A., S. Sendra, J. M. Jimenez et J. Lloret. 2019, «Dynamic metric ospf-based routing protocol for software defined networks», *Cluster Computing*, vol. 22, n° 3, p. 705–720. 78
- RY. Rubinstein. May 1997, «Optimization of Computer Simulation Models with Rare Events», *European Journal of Operational Research*, vol. Volume 99, n° 1, p. p. 89–112, ISSN 0377-2217. 85
- Spiteri, K., R. Urgaonkar et R. K. Sitaraman. 2016, «Bola : Near-optimal bitrate adaptation for online videos», dans *The 35th Annual IEEE International Conference on Computer Communications (INFOCOM'16)*, p. 1–9. 84, 88
- Serritt, R. 2004, «Autonomic networks : engineering the self-healing property», *Engineering Applications of Artificial Intelligence*, vol. 17, n° 7, doi :<https://doi.org/10.1016/j.engappai.2004.08.028>, p. 727–739, ISSN 0952-1976. URL <https://www.sciencedirect.com/science/article/pii/S0952197604001113>, autonomic Computing Systems. 76
- Sun, Y., X. Yin, J. Jiang, V. Sekar, F. Lin, N. Wang, T. Liu et B. Sinopoli. 2016, «Cs2p : Improving video bitrate selection and adaptation with data-driven throughput prediction», dans *Proceedings of the 2016 ACM SIGCOMM Conference (SIGCOMM'16)*, ACM, p. 272–285. 84
- Temlyakov, V. 2015, «Greedy approximation in convex optimization», *Constructive Approximation*, vol. 41, n° 2, p. 269–296. 81

- Tong, V. February 2021, «Service-centric segment routing using reinforcement learning», URL <https://github.com/vanvantong/rl-sr>. 82
- Tong, V., H. A. Tran, S. Souihi et A. Mellouk. 2020, «Service-centric segment routing mechanism using reinforcement learning for encrypted traffic», dans *2020 16th International Conference on Network and Service Management (CNSM)*, IEEE. 78
- Tran, H. A., S. Hoceini, A. Mellouk, J. Perez et S. Zeadally. 2013, «Qoe-based server selection for content distribution networks», *IEEE Transactions on Computers*, vol. 63, n° 11, p. 2803–2815. 81
- T.Y Huang, N. M. M. T., R. Johari et M. Watson. August 2014, «A buffer-based approach to rate adaptation : Evidence from a large video streaming service», dans *Proceedings of the 2014 ACM Conference on SIGCOMM (SIGCOMM'14)*, ISBN 978-1-4503-2836-4, p. 187–198. 84, 88
- V. Krishnamoorthi, E. H., N. Carlsson et E. Petajan. June 2017, «Buffest : Predicting buffer conditions and real-time requirements of http(s) adaptive streaming clients», dans *ACM SIGMM Conference on Multimedia Systems (MMSys'17)*, p. 1–12. 84
- Yap, K.-K., M. Motiwala, J. Rahe, S. Padgett, M. Holliman, G. Baldus, M. Hines, T. Kim, A. Narayanan, A. Jain et collab.. 2017, «Taking the edge off with espresso : Scale, reliability and programmability for global internet peering», dans *Proceedings of the Conference of the ACM Special Interest Group on Data Communication*, p. 432–445. 78
- Zahran, A. H., J. Quinlan, D. Raca, C. J. Sreenan, E. Halepovic, R. K. Sinha, R. Jana et V. Gopalakrishnan. May 2016, «Oscar : An optimized stall-cautious adaptive bitrate streaming algorithm for mobile networks», dans *Proceedings of the 8th International Workshop on Mobile Video (MoVid'16)*, p. 1–6. 84, 86, 88
- Zhou, C., C. W. Lin, X. Zhang et Z. Guo. 2013, «Buffer-based smooth rate adaptation for dynamic http streaming», dans *2013 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference*, p. 1–9. 84, 88

Chapitre 6

- Almadani, B., A. Beg et A. S. H. Mahmoud. 2021, «Dsf : A distributed sdn control plane framework for the east/west interface», *IEEE Access*, vol. 9, p. 26 735–26 754. 98
- Aujla, G. S., M. Singh, A. Bose, N. Kumar, G. Han et R. Buyya. 2020, «Blocksdn : Blockchain as a service for software defined networking in smart city applications», *IEEE Network*, vol. 34, n° 2, p. 83–91. 100
- Bailey, J. et S. Stuart. 2016, «Faucet : Deploying sdn in the enterprise», *Commun. ACM*, vol. 60, n° 1, doi :10.1145/3009828, p. 45–49, ISSN 0001-0782. URL <https://doi.org/10.1145/3009828>. 99
- Bannour, F., S. Souihi et A. Mellouk. 2017a, «Distributed sdn control : Survey, taxonomy, and challenges», *IEEE Communications Surveys & Tutorials*, vol. 20, n° 1, p. 333–354. 98
- Bannour, F., S. Souihi et A. Mellouk. 2017b, «Software-defined networking : A self-adaptive consistency model for distributed sdn controllers», *RESCOM 2017*. 98

- Berde, P., M. Gerola, J. Hart, Y. Higuchi, M. Kobayashi, T. Koide, B. Lantz, B. O'Connor, P. Radoslavov, W. Snow et G. Parulkar. 2014, «Onos : Towards an open, distributed sdn os», dans *Proceedings of the Third Workshop on Hot Topics in Software Defined Networking*, HotSDN '14, Association for Computing Machinery, New York, NY, USA, ISBN 9781450329897, p. 1–6, doi :10.1145/2620728.2620744. URL <https://doi.org/10.1145/2620728.2620744>. 99
- Brockelsby, W. et R. Dutta. 2021, «Traffic analysis in support of hybrid sdn campus architectures for enhanced cybersecurity», dans *2021 24th Conference on Innovation in Clouds, Internet and Networks and Workshops (ICIN)*, p. 41–48, doi :10.1109/ICIN51074.2021.9385530. 98
- Gerola, M., F. Lucrezia, M. Santuari, E. Salvadori, P. L. Ventre, S. Salsano et M. Campanella. 2016, «Icona : A peer-to-peer approach for software defined wide area networks using onos», *2016 Fifth European Workshop on Software-Defined Networks (EWSDN)*, p. 37–42. 98
- Jiasi, W., W. Jian, L. Jia-Nan et Z. Yue. 2019, «Secure software-defined networking based on blockchain», *arXiv*. 100
- May, L. et T. Runyon. 2019, «Labscrum : A case study for agility in academic research labs», . 100
- Medved, J., R. Varga, A. Tkacik et K. Gray. 2014, «Opendaylight : Towards a model-driven sdn controller architecture», *Proceeding of IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks 2014*, p. 1–6. 99
- Nam Nguyen, H., H. Anh Tran, S. Fowler et S. Souihi. 2021, «A survey of blockchain technologies applied to software-defined networking : Research challenges and solutions», *IET Wireless Sensor Systems*, vol. 11, n° 6, p. 233–247. 100
- Schwaber, K. et J. Sutherland. 2017, «The scrum guide™», . 100
- Yu, H., H. Qi et K. Li. 2020, «Wecan : an efficient west-east control associated network for large-scale sdn systems», *MONET*, vol. 25, n° 1, doi :10.1007/s11036-018-1194-9, p. 114–124. URL <https://doi.org/10.1007/s11036-018-1194-9>. 98